

2000

A comparison of the use of modal verbs in research articles by professionals and non-native speaking graduate students

Jenny Marie Hykes
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>



Part of the [Bilingual, Multilingual, and Multicultural Education Commons](#), [English Language and Literature Commons](#), and the [First and Second Language Acquisition Commons](#)

Recommended Citation

Hykes, Jenny Marie, "A comparison of the use of modal verbs in research articles by professionals and non-native speaking graduate students" (2000). *Retrospective Theses and Dissertations*. 7929.
<https://lib.dr.iastate.edu/rtd/7929>

This Thesis is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.

A comparison of the use of modal verbs in research articles
by professionals and non-native speaking graduate students

by

Jenny Marie Hykes

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
MASTER OF ARTS

Major: English (Teaching English as a Second Language/Applied Linguistics)

Major Professor: Susan Conrad

Iowa State University

Ames, Iowa

2000

Graduate College
Iowa State University

This is to certify that the Master's thesis of

Jenny Marie Hykes

has met the thesis requirement of Iowa State University

Signatures have been redacted for privacy

TABLE OF CONTENTS

ABSTRACT	vi
CHAPTER 1. INTRODUCTION	1
Importance of this Study	2
Research Questions	4
Preview of the Study	5
CHAPTER 2. LITERATURE REVIEW	6
Functions and Frequencies of Modal Verbs	6
Forms of modals	7
Semantics of modals	8
Extrinsic and intrinsic modality	8
Ambiguous meanings	9
Frequencies in different registers	10
Use of Modal Verbs in Scientific and Research Writing	11
Reasons for modality in scientific writing	11
Frequencies of modal verbs in scientific corpora	13
Frequencies in research article sections	14
Functions of modal verbs in research articles	16
Summary of modality in scientific writing	17
Nonnative Speaker Use	18
Challenges of modal verbs	18
Textbook studies	19
Studies of modal verbs in NNS writing	20
Methodology of Corpus Linguistics for Interlanguage	23
Principled corpora	23
Learner characteristics	24

Native speaker characteristics	24
Characteristics of the writing	25
Types of Analysis	26
Summary	27
 CHAPTER 3. METHODS	 28
The Learner Corpus	28
Characteristics of the students	28
Characteristics of the writing	29
The Professional Corpus	31
The writers	31
The texts	32
Quantitative Analysis	33
Qualitative Analysis	34
Summary	35
 CHAPTER 4. RESULTS AND DISCUSSION	 36
Quantitative Findings	36
Overall counts	36
Range in overall counts	38
Research article section counts	39
Ranges in section counts	41
Counts of individual modals	42
Occurrences of <i>may</i>	42
Occurrences of <i>can</i>	43
Occurrences of <i>will</i>	44
Functions of the Modal Auxiliaries in the Corpus	45
An overview of the semantic categories of <i>may</i> and <i>can</i>	45
Epistemic use of <i>may</i> most common	48
Semantic classifications of <i>can</i>	52

Student use of <i>can</i>	53
<i>Can</i> substituted for <i>may</i>	54
<i>Can</i> more often intrinsic	54
<i>Can</i> used to express legitimacy	55
Ambiguous use of <i>can</i>	57
Use of <i>will</i>	60
Student use of <i>will</i> in Introduction sections	61
<i>Will</i> indicating volition	61
Making predictions	62
Student use of <i>will</i> in Methods sections	65
Professional use of <i>will</i> in Discussion sections	66
Summary: Professional-Student Differences in Modal Use	68
 CHAPTER 5. CONCLUSION	 70
Major Findings	70
Implications for Materials Development and Current Instruction	73
Implications for Further Research	78
Conclusion	78
 APPENDIX A: CONSENT FORM	 80
 APPENDIX B: HUMAN SUBJECTS FORM	 82
 WORKS CITED	 86
 ACKNOWLEDGEMENTS	 89

ABSTRACT

This study compares the use of modal verbs in research articles by international graduate students compared with use by professional writers of scientific research articles. To make these comparisons, two corpora of research article texts--one by international students and one by professionals--were created and modal verbs were counted in each of the corpora. Total frequencies of modal verbs, frequencies in each research article section and frequencies of individual modals were all compared. The functions for which students and professionals used particular modal verbs were also compared.

The findings showed that students used modal verbs with twice the frequency of their professional counterparts throughout the corpora. The modal verbs *can* and *will* were particularly overused by students, while *may* was the one modal consistently underused by students. Students often used these modals for functions uncommon in professional use.

These findings will be useful to ESL teachers and materials writers who are interested in helping students write in a more professional, effective style. This research also contains many applications for further research into learner modal verb use in scientific research articles.

CHAPTER 1. INTRODUCTION

In presenting their findings to the scientific community, scientists must make myriad choices regarding how to reduce months of research, usually with plenty of human ups and downs, setbacks and surprises, into a few pages of scientific knowledge. Included in these choices are questions of how to present their most interesting, unexpected or important findings as specifically and carefully as possible, in a way that discreetly acknowledges the limited nature of all research, but also conveys the importance of their own work (Swales, 1990). Modal auxiliary verbs, such as *may*, *can*, *might* and *could*, become essential tools in achieving that balance between new exciting claim and carefully expressed proposition.

When professionals who are native speakers of English make their choices about how to weigh in on this balancing act, they come with a far greater advantage than their non-native speaking counterparts. Intuitively these native speakers comprehend that modal verbs do not serve one sole and simple purpose within English, or even within the specific genre of scientific writing. They realize any one of these forms has more than one possible meaning and function, including permission, ability or possibility, obligation or necessity, volition or prediction. As they understand what functions are appropriate in which context, they are able to choose the appropriate forms. And they understand the subtle gradations between certainty and doubt conveyed in the forms expressing possibility. None of this knowledge comes easily or naturally to the non-native speaker, and yet, it is

essential for them to handle these forms with the same precision that native speaking scientists do, if they wish for their tentative findings to be expressed accurately.

In order to discern how the use of modal verbs may be interfering with the attempts of non-native speaking international graduate students to make clear and qualified claims and to represent themselves as competent members of the scientific discourse community, this study compares the use of modal verbs in scientific research articles by international students and professional, published scientists.

Importance of this Study

This study's importance lies in its focus on one form, modal verbs, and their use by non-native speakers in the specific genre of research articles. Information about professional and learner use of modal verbs in research articles may assist the learners, their teachers and textbook writers to understand more about the appropriate or common frequencies and functions of modal verb use in professional research articles, and the ways students vary from these accepted uses. ESL students and teachers will also benefit from considering the implications of the findings for student needs in research writing.

The focus on modal verbs is significant because of their complexity in both syntactical form and semantic meaning, which makes them challenging for non-native speakers to learn and use in any setting, and because of their importance in all communication, including the particular type of communication important for international graduate students, scientific research articles (Holmes, 1983, 1988;

Stubbs, 1985).

The research article is a form most international students must master in order to proceed in academic and research careers; scientists who do research need to publish their findings in order to fully participate in the discourse of their community. Not only is the research article important in their future careers, but also in their graduate coursework, in which professors' will expect them to write up their findings for class projects in research article format.

Expressions of modality are important in scientific writing as a way to present new claims with appropriate politeness and precision (Holmes, 1988; Hyland, 1996; Myers, 1989). Although many forms express modality, including modal adverbs, adjectives and lexical verbs and nouns, limiting this study to modal verbs allows me to look closely at the use of the most common surface form used to express modality (Biber et al, 1999; Holmes, 1988; Stubbs, 1985).

The choice to study non-native speaker use is important because of the potential to offer useful information to learners and their teachers. While past research has been done regarding the use of modal verbs and other forms of modality in the scientific writing of native speakers (Adams Smith, 1984; Butler, 1990; Hyland, 1996; Salager-Meyer, 1994; Skelton, 1997; Varttala, 1998), I found no research regarding the use of modal verbs by non-native speakers in this genre. Flowerdew (1997) looked at undergraduate students' writing of reports based on first-hand data, including surveys, interviews and observation. While useful, her study did not provide much information about the more complex and more important writing of scientific research articles for potential publication in international journals.

Thus this study is important for three related reasons. First, it looks closely at only one form, allowing for more understanding of that particular form in its complexity. Second, it focuses on the specific and important genre of scientific research articles. Finally, it gives an understanding of non-native speaker use in that genre, an area of research that has not yet been studied.

Research Questions

The purpose of this study was to determine whether NNS graduate students use modal verbs differently than professionals when writing scientific research articles. Modal verb use in a corpus of research articles from both groups was examined. The study looked at the contexts and functional purposes that influence the choice to use modal verbs. Comparisons between the two groups were made to determine how and why learners of English use modal verbs differently than professionals, as well as what effects such differences have on NNS writing.

To fulfill this purpose, the study addressed the following research questions:

- How does the use of modal verbs in scientific research articles differ between NS professionals and advanced international graduate students with respect to
 - the overall frequency of modal verbs used?
 - the frequency differences between research article sections?
 - frequency of individual modals used (overall and by sections)?
 - the functions of the modal verbs used?

Preview of the Study

To address these research questions, I begin in Chapter 2, Literature Review, with a presentation of background information about modal verbs and their frequencies in different genres and the past research of modal verbs in scientific writing and in the writing of non-native speakers of English. In Chapter 3, Methods, I explain the two corpora created for the study, one of international students' writing and one of writing by professionals, and the quantitative and qualitative analyses done of the modal verbs that appeared in these two corpora. The results of those analyses are presented and discussed in Chapter 4. Finally, in Chapter 5, I summarize these findings and discuss major suggestions for further research and implications for learner and their teachers.

CHAPTER 2. LITERATURE REVIEW

Studying the differences in modal verb use requires background in several areas: the general frequencies of modal verbs across registers and between different registers; the use of modal verbs and other forms of modality in writing, particularly scientific and research writing; and the use of modal verbs by nonnative speakers. Because the present study uses corpus linguistics, I also need sufficient background in the use of corpus linguistics for studying interlanguage. This chapter, therefore, introduces the defining characteristics of modal verbs, their typical semantic meanings and their functions and frequencies in different registers. It discusses the specific findings regarding modal verb use in scientific writing, including the most common modal verbs, their typical functions and their varied use throughout the different sections of the research article. Finally, this chapter introduces the key requirements of corpus linguistics.

Functions and Frequency of Modal Verbs

Any study of modal verb use requires first a definition and list of modal verbs, their grammatical characteristics, semantic meanings, general functions and relative frequencies in different contexts. However, a complete discussion of even merely the semantics of modal auxiliaries is far beyond the scope of this paper. Here I give an overview of the forms, meanings and functions of modal verbs, and then discuss the frequencies of these forms within different registers, an area of greater relevance

to my own research.

Forms of modals

The modal auxiliaries are generally considered to be the verbs *can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will*, and *would* (Biber et al., 1999). These verb forms are distinguished from lexical verbs by the following characteristics (Halliday, 1976):

1. They have no finite, *to*, *-ing*, *-ed*, etc. forms: *"*maying*, *to may*."
2. Negative and interrogative forms are made without expansion. It is not necessary to include a "to do" or "to be" form when creating questions or making negative statements with modal verbs. "He will not." "Can she go?" rather than *"*Is she can go?*" and *"*He doesn't will.*"
3. Negative forms are reducible. They can be contracted, as in "I can't write this paper." "I won't be able to finish."
4. They can be used as "code verbs" in ellipsis: This means that when two equivalent clauses using a modal verb occur together, the second clause does not need to repeat the accompanying lexical verb, as in, "I can go. So can she." In this case the second modal verb (*can*) represents the lexical verb (*go*).
5. They do not require a third person singular *-s*: *"*She cans.*"
6. They are not found in imperative clauses.
7. They do not combine with each other and do not co-occur in a clause (except in some nonstandard dialects: "She *might could* go.").

Despite these clear characteristics, Biber et al. (1999) argue that the distinction between modal verbs and lexical verbs is not completely clear. Some normally non-modal lexical verbs can also function as modal verbs, especially in spoken contexts. These forms are classified in two categories. One category is marginal auxiliary verbs and includes *need to*, *dare to*, and *used to*. These verbs have features which are characteristic of modal verbs, such as similar negative and interrogative forms: "*needn't*, *dare not*, *need we?*." Biber et al. (1999) also include *ought to* in this group. The other category is called semi-modals or quasi-modals: *(had) better*, *have to*, *(have) got to*, *be supposed to*, *be going to*. These idiomatic expressions act as forms of modality but can also be marked for tense and person and can co-occur with modal verbs (Biber et al., 1999).

This study's focus, based on the information from the current literature regarding the defining characteristics of modal verbs, is on the following verbs: *can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will*, *would*, *need to*, *dare to*, *ought to*, *(had) better*, *have to*, *(have) got to*, *be supposed to*, *used to* and *be going to*.

Semantics of modals

Extrinsic and intrinsic modality

While the list of relevant verbs is fairly limited, the list of possible semantic functions these verbs fulfill is rather extensive. There is not a one-to-one correspondence of meaning and form, and most modal verbs can fill more than one semantic function. Two basic functional distinctions, accepted by most linguists, are

made between extrinsic and intrinsic modality (also known as "epistemic" and "root" modality) (Biber et al., 1999). Most modal verbs can be either extrinsic or intrinsic, depending on the functions they fill and the meanings they relate.

Extrinsic modality is concerned with the "logical status of states and events" (Biber et al., 1999), and is based on the reality outside of the inner reality of the text. These extrinsic meanings (commonly called epistemic modals) are often related to the subjective viewpoint of the speaker with regard to doubt or certainty of a proposition (Palmer, 1979). These verbs suggest likelihood, possibility or prediction of a proposition (Biber et al., 1999). For example: "The weather *may* turn bad this evening."

Intrinsic modality speaks about agent-controlled actions and events. These verbs express obligation, volition, permission, and ability (Biber et al., 1999; Palmer, 1979). For example: "I *will* chose the yellow pencil." "You *ought* to go to be earlier." "This computer *can* do many things."

These characteristic divisions according to semantics provide a framework for analysis of the functions of modal auxiliaries. While frequency of occurrences is one telling aspect of use, the differences in intrinsic and extrinsic use are also important.

Ambiguous Meanings

The difficulty of assigning rigid semantic meanings to the modal auxiliaries was discussed by Leech and Coates (1980) and Coates (1983). They argued against using logical formalism to set up definitions of the words and suggested instead an understanding of the "indeterminacy" of meanings in these words. Three

types of indeterminacy--gradience, ambiguity, and merger--can help distinguish between words that fit somewhere between two categories, or that can have two separate meanings and can logically fit in one or both of these semantic categories (Leech and Coates, 1980; Coates, 1983). This ambiguity of semantic classification suggests that analysis of modal verb use cannot be expected to neatly divide all modal verbs into only two rigid categories.

Frequencies in different registers

Past research has found that modal verbs with all their varied meanings occur to different degrees throughout various discourse registers. The term "register" is used to mean varieties of discourse "defined by their situational characteristics" (Biber, Conrad, & Reppen, 1998). Registers can be more general and broad, such as registers of conversation and writing, or more specific, such as registers of newspaper writing and academic writing, or even more specialized, such as the register of editorials and the register of sports reporting, the register of composition essays and the register of research articles. In Holmes' (1983) study modal verbs occur more often in the register of speech and especially in conversation (42.4 per 1000 words) than in writing (36.8 per 1000 words). Comparing conversation, newspapers, fiction and academic prose, Biber et al. (1999) found that modal and semi-modal verbs were used most often in conversation (25 occurrences per 1,000 words) and least often in news writing and academic prose (14 occurrences per 1,000 words in academic prose). While the numbers in these two studies differ, both point to a greater use of modals in conversation than in writing. The previous

research shows, then, that modal verbs are a more common feature of oral discourse, and appear least frequently in academic writing.

Use of Modal Verbs in Scientific and Research Writing

Although modal verbs occur more often in oral communication, the preceding information indicated that modal verbs play a crucial role in writing. The role of epistemic modality in academic writing specifically has been of recent interest to applied linguists. Researchers have begun discussing the important function of epistemic forms of modality (also referred to as "author comment") in even the most scientific and objective writing. Some studies covered the broader topic of modal forms which were not limited to modal verbs, while others were focused only on modal verbs.

Reasons for modality in scientific writing

Forms of modality, including modal verbs, have been shown to serve at least four related functions in scientific writing: to express politeness; to suggest tentativeness; to mark a proposition as the writer's own viewpoint; and to create a scientific atmosphere. These different functions often occur simultaneously, and we see in all of the functions that forms of modality are used in academic and specifically scientific writing to show the place of the writer in relationship to both the text and the reader. In some cases forms of modality serve as politeness strategies, achieving politeness through mitigations, often called "hedges" (Hyland, 1996;

Myers, 1989; Salager-Meyer, 1994; Varttala, 1998). This politeness is required because of the particular nature of scientific discourse, which requires researchers to present new findings in an extremely accurate manner which also acknowledges the contributions of other scientists. Varttala's research (1998) also suggested hedges are used in more general, popular scientific journals to create a scientific atmosphere and prove the credibility of the writings.

In other cases modal verbs indicate the author's attitude towards any statement within the text (Adams Smith, 1984), and show the writer's role as the "understood agent," someone who is reporting his or her own views (Thompson and Yiyun, 375:1991).

Scientists also employ modal verbs in order to present their claims more accurately. Expressions of modality

allow writers to express propositions with greater precision, recognizing the impossibility of exactly quantifying the world. Scientific writers seek to balance fact and interpretation as they try to present information as accurately as possible, always recognizing that only a segment of reality can be described... (Hyland, 478:1996).

Hedges are also used to suggest the appropriate degree of tentativeness which would allow them to acknowledge the possibility that they could be proven wrong and the fact that they are only expressing their perspectives about their findings rather than certain truth (Adams Smith, 1984; Butler, 1990; Hyland, 1996; Myers, 1989; Varttala, 1998).

Modality in scientific writing is expressed through several different forms: lexical verbs, adverbs and adjectives, modal nouns and of course, modal verbs. As Butler

says, "The modal verbs are one of the most powerful devices available in English for the presentation of conclusions with a range of subtle gradations in strength and confidence" (Butler, 1990). Because of the complexities and importance of modal verbs in scientific writing, this study focuses on the ways learners of English handle these verbs to function as forms of politeness and textual qualification, as well as for other purposes.

Frequencies of modal verbs in scientific corpora

Several researchers have found a significant number of modal verbs in corpora of scientific writing. Butler's (1990) research found modal verbs made up 1 percent of all words in his corpus, while Hyland (1996) noted that a limited number and variety of modal verbs functioning as hedges made up his scientific corpus. Modal verbs made up half of all the author comments appearing in a corpus of medical writing in Adams Smith's study (1984).

These researchers have also found that certain modal verbs were more likely to appear in scientific writing. Among these, *may*, *would*, *could* (Hyland, 1996), *may* and *should* (in medical writing (Adams Smith, 1984)), and *may* and *can* (in physics, botany and plant physiology writing (Butler, 1990)) have been found to be more preferred in scientific writing than in other forms of writing. Butler (1990) found the auxiliaries *will*, *would*, *could* and *should* were less appropriate in scientific writing than in more general genres of writing. Hyland's (1996) study also suggested certain hedging forms--including *may*, *could*, and *should*--are more common in scientific writing than in other forms of writing.

The use of modal verbs in scientific writing has shown variation based on topics, audiences, and research article sections. For example, Butler (1990) found that biology texts were twice as likely to have occurrences of *may* and *might* as physics texts, and that *can* appeared more often in physics texts. He suggests that the differences relate to topic. The literature suggests the importance of modal verbs within the context of scientific writing, and implies the need for greater research into the questions of learner use of modal verbs in scientific writing.

Frequencies in research article sections

The majority of research into the varieties of modal verb use has been in the area of research article sections. Skelton looked at the "types of truth" in medical writing, as a way to explain the differences in modal use within scientific research articles (1997). He and others found that the "contextualized truth," which occurs primarily in Introductions for discussing the research of others does not require frequent occurrences of epistemic modality (Skelton, 1997 and Thompson and Yiyun 1991). In Results sections and to some extent in Discussion sections, where "evidential truth" based on concrete evidence--numbers and statistical significance--is important, verbs like *show* are more common and there are fewer instances of modality (Skelton, 1997). What Skelton labeled "interpretive truth" is the most important context for modal verbs. This context, which occurs in some Results sections and even more often in Discussion and Conclusion sections, requires a greater reliance on modality expressions because of its characteristic of relating non-evidential claims and judgements of truth and doubt and certainty. These

sections often deal with what *may* be the case and what *should* be the case. Discussion sections also see a larger proportion of modality because of their tendency to offer suggested action and future research (Skelton, 1997).

Other research backs up Skelton's hypothesis. Varttala found the modal auxiliaries *may*, *might* and *could* used in their epistemic meaning as hedging devices occur most frequently in the Discussion sections of research articles appearing in medical journals, and occur only half as much in Introductions, even less in Results sections, and were almost non-existent in Methods sections (1998). General studies found that author comments and hedging occurred more often in Discussion sections (Adams Smith, 1984; Salager-Meyer, 1994). Similarly, the specific studies of modal verbs by Butler and Varttala found modal auxiliaries occurred more often in Discussion sections (Butler, 1990; Varttala, 1998). This increase in expressions of modality reflects the "switch from detachment to involvement" (Adams Smith, 28: 1984). In the Discussion sections, there was one occurrence of author comment for every 2.2 lines, and more than half of these occurrences were modal auxiliaries (primarily *may* and then *should*), with most cases of *should* appearing in a concluding sentence (Adams Smith, 1984). Similarly, Salager-Meyer found that 13% of the words in Discussion sections were hedges using modals to qualify or enhance quantitative and qualitative findings (1994). She explained these findings:

It is in this last section of research papers that writers speculate, argue, contrast and extrapolate from the described results, and at the same time avoid stating result too conclusively so that the readers can note that the authors are not claiming to have the final word on the subject.... By resorting to such expressions, researchers can avoid absolutes and thus indicate exactly the degree of certainty with which they present their

conclusions and also how strongly they want to align themselves with their claims (Salager-Meyer, 1994).

Researchers general consensus about the proportions of modality typically used in the other sections of a research article suggests concentrations are found most often in Discussion sections, then Introductions, then Results and finally, least frequently in Methods sections (Adams Smith, 1984; Butler, 1990; Salager-Meyer, 1994). This fairly universal finding of the relative proportions of modal verbs in the different research article sections indicates a study of modal verb use needs to include the differences of use across article sections.

Functions of modal verbs in research articles

Most of this past research looked at the ways expressions of modality are used to qualify writers' statements, and so concentrated on instances of epistemic modality as opposed to modality performing any other function. In a study of the various functions of author's comments in medical research writing, Adams Smith (1984) found that author's comments served the following functions, in order of frequency: 1. Assessing probability or possibility related to the truth or definiteness of a thesis. 2. Recommending or expressing obligation or necessity. 3. Evaluating. 4. Emphasizing comments. 5. Expressing ability. 6. Arguing for a point or action. 7. Expressing expected or unexpected outcomes. Salager-Meyer (1994) also found that epistemic functions of modal verbs were the most common in research articles, most often used to express tentativeness and to qualify statements, and occasionally, (in Methods sections) to act as approximators.

Butler (1990) looked specifically at the functions of certain auxiliary verbs and found *may* and *might* were frequently epistemic while *can* is less likely to be epistemic. His corpus had only one occurrence of an epistemic *can*. Most instances of *will* referred to future time, and some were used to show inanimate ability (intrinsic), or a deductive relationship between two statements. Only one example of *will* served as pure prediction (extrinsic). Only two uses of *would* were found in expressions of tentativeness, co-occurring with verbs such as *seem*, *appear* and *expect*. And *should*, *shall* and *must* were all more likely to function as expressions of obligation, although *must* also did serve epistemic/extrinsic functions.

This past research, in particular Butler's, which was the only study to look specifically at modal verb use in scientific research articles, suggests the importance of the differing functions of modal verbs within different contexts. Whether NNS student writing follows the conventional uses remains an important research issue.

Summary of modality in scientific writing

Expression of modality is an important aspect of scientific writing, and modal verbs, which are the most common forms of modal expression, are perhaps the most important forms of modality. They serve to qualify statements and are crucial in presenting the scientific researchers' views of their findings in a manner suggesting both politeness to the larger scientific community and the specific differences between research findings and unequivocal facts. They are more likely to be found in the Discussion sections of research articles, and are least common in Methods sections. The verbs *may* and *can* are especially common in scientific writing,

although other forms also occur. *May* is more likely to be used in its extrinsic functions, while *can* is more likely to be used intrinsically.

Nonnative Speaker Use

Challenges of modal verbs

As useful as modal verbs are in communication, including academic and scientific writing, they are not easy to learn or to use appropriately for nonnative speakers of English (Gibbs, 1990). English L2 learners may have a relatively simple time learning the surface positions of modal verbs but may have a bigger problem recognizing and using modal verbs properly with respect to their underlying meaning (Cook, 1978). There are a few possible reasons for this trouble. One problem with modal verbs, and all expressions of modality, is that the linguistic forms do not have a one to one relationship with meaning (Holmes, 1988). Furthermore, categories involved in expressing modal meaning are not clear, but rather are better understood as “degrees of certainty or epistemic commitment to the validity of a proposition.” (Holmes, 1982; Leech G. and Coates J., 1980). Another challenge for NNS is the fact that the use of modality, like other pragmatic features, is culturally determined (Hinkel, 1995; Holmes, 1988). Since epistemic modality is used both to express the speaker’s perspective about a proposition, as well as deference to one’s addressee, NNS students need to develop sociocultural sensitivity to learn to use it appropriately (Holmes, 1983). Not only do learners struggle with using modal forms, but also with recognizing their meaning and range of meanings in reading, causing confusion

between accepted facts and objective statements, especially in scientific writing (Adams Smith, 1984).

The research suggests students may have difficulty controlling their use of modal verbs, especially within the context of the complicated and demanding task of writing up their own research in a research article.

Textbook studies

Studies of textbooks (Holmes, 1988; Hyland, 1994) indicate that students may receive inadequate information about modality and modal verb use. Course books studied by Holmes (1988) used less than half (40% and 27%) of the total range of epistemic modal verbs in their content and grammar exercises, did not present other forms of modality other than auxiliary verbs, and did not present information about the frequencies of use of these different forms. Hyland's (1994) study of EAP and ESP textbooks noticed that few texts presented information about the extrinsic use of modal verbs or only covered *would*, *should*, *can* and *may* and even in these cases did not make a clear distinction between extrinsic and intrinsic modality. Sometimes the textbook presentation of modal forms is misleading because it has been oversimplified. In one textbook Holmes investigated, modal verbs were described as "impersonal, implying that assessment of their validity is based on some objective measure." In another text modal verbs and other epistemic expressions were only discussed in their function as expressions of affective meanings of politeness, rather than modal meanings (Holmes, 26:1988).

After looking at these problems, Holmes argued that it is important for learners

to be taught a range of possible expressions of doubt, uncertainty, possibility and probability and "some awareness of the relative frequency of different devices in different contexts, so that they can accurately gauge the stylistic effect of using one rather than another" (Holmes, 23:1988).

If students are not taught that there are more forms of modality other than modal verbs commonly used by writers, such as modal adverbs and adjectives, and lexical nouns and verbs, they may overuse modal verbs in their scientific writing. If they are not taught the nuances of meanings and appropriate uses for modal verbs within certain contexts, the difficulties of controlling use may be even more apparent.

Studies of modal verbs in NNS writing

Studies of NNS writing have demonstrated problems with over and underuse of modal verbs. More generally, many studies have suggested learners are more likely to use a more "direct" style of writing than NS without the qualification or mitigation that forms like modal verbs supply (Flowerdew, 1997, 1998; Hu et al., 1982; Hyland and Milton, 1997).

Hyland and Milton (1997) found that the L2 writing was more likely to convey a stronger degree of certainty, with "firmer assertions, a more authoritative tone and strong writer commitments," when compared with NS writing (193). Their results showed NNS students used 60% more certainty markers than L1 students, while the NS students used 73% more expressions of probability. Similarly, half the epistemic expressions used by the Chinese students functioned to show the writers' certainty while epistemic markers in the NS writing functioned to qualify the probability or

doubt of an assertion. In this study, while both groups made extensive use of modal verbs to make a qualification, NNS actually overused most modal verbs (they used *will* and *may* twice as often as NS students, but *would* and *might* with half the frequency) and most lexical verbs expressing doubt and certainty, but they used adverbials as expressions of modality to a much lesser extent than native speakers. The researchers also noticed that the Chinese students who scored the highest grades on the high school graduation exam used the least modal verbs.

Learners used modal verbs infrequently, especially in epistemic or extrinsic function, in Flowerdew's (1997) study of a learner corpus of analytical writing by 2nd and 3rd year university students in Hong Kong. The resulting writing often was "too direct" for the context, in Flowerdew's estimation (79:1997). Offering anecdotal evidence, she concluded that students' interpretation of results is overly certain, based on "certain knowledge, rather than plausible knowledge," (80:1997). She suggested students have a limited lexicon for commenting on and reporting data.

In Hu, Brown, and Brown's (1982) analyses of the interpersonal metafunction of Chinese and Australian students' writing in English, it was found that Chinese students preferred "root" or intrinsic modals, and semi-modals, such as "have to" and "be able to" over extrinsic modals expressions of probability and predictability, which were more often used by the Australian students. The Chinese writers used the expressions "If...not, you'll...", as in "If you don't work hard now, you'll regret it when you grow up," with a greater degree of frequency, suggesting to the researchers that these writers had a more definitive way of thinking than the Australian writers. Overall, the Australian students were twice as likely to use expressions of modality

as Chinese students.

In Flowerdew's (1998) study of learners' expressions of causality, she found the professional, native speaker corpus showed a greater tendency to modify causal verbs with modal verbs than the corpus of writing by Hong Kong university students. She also found that learners were less likely to use mitigating devices, especially modal verbs, with causative phrases. Learners almost never used *can* and *may* in these environments, while NS often used modal verbs and adverbs in these cases. Flowerdew suggested these findings indicate that learners attempt to show direct causal relationships when less definite relationships may be the reality. Another reason for the findings, I believe, may be because of learners' difficulties with expressions of modality, including modal verbs.

Hinkel (1995) considered the problem of the cultural context implicit in modal verb use in her study. She found that ESL students from Asian cultures with a Confucian/Taoist influence used modal verbs of obligation and necessity differently from the way NS used those same forms when writing about issues concerning family, tradition, education, and patriotism. She suggested the reason for this was their differing values of what is obligation and necessity. Hinkel argued that students need to learn modals in L2 contexts and points of reference to enable them to use them similarly to NSs. This argument may be especially relevant when learners are writing scientific and academic texts and wish to be heard by a certain discourse community. In these cases learners need to conform to the rhetorical standards of that community's culture, especially the standards of politeness and deference. Discrepancies between the way NNS and NS use modal verbs in these kinds of

contexts would be more problematic.

The findings do not present a clear consensus about typical student use of modal verbs. In some cases, it seems overuse is the prevailing problem, causing writing to sound too conversational and unsophisticated, while in other studies, students' underuse of the verbs causes their writing to sound more direct than native speaker writing. Furthermore, no studies have examined a large number of advanced students writing research papers.

Methodology of Corpus Linguistics for Interlanguage

Principled corpora

Because this is a corpus linguistic study of interlanguage, the requirements of a well-designed corpus study must be considered. Biber, Conrad & Reppen (1998) say a principled corpus should be of an adequate size (number of words, texts and registers) large enough to be representative and generalizable, and a diverse and well-reasoned array of source (i.e.: not all by the same writer, not all about the same subject). In most studies of learner language, including my own, two separate corpora are compared usually in order to study differences in frequencies of use of a certain form or forms. These types of studies require two principled corpora that differ as little as possible, except in the characteristic under investigation. Creating a learner corpus and finding an appropriately comparable NS corpora are some of the biggest challenges for researchers in interlanguage corpus linguistics, and this challenge is no less pronounced in my own situation.

Past researchers have had to consider several features when devising an acceptable corpus, including level of the students and their L1, the level of the NS corpus: students or professionals, and the writing topics, registers and task characteristics of the writing of both corpora.

Learner characteristics

The majority of interlanguage studies have focused on writing done by students of English in countries where English is not a primary language, either for reasons of necessity (the researchers are in an EFL context), or to avoid the effects of immersion experiences on the learner language. Generally, the students have been at the university level, although a few studies have used students graduating from high school. All the students whose writing makes up a corpus are typically at a uniform level. Sylviane Granger, the creator of the most prevalent corpus of learner language, the International Corpus of Learner English (ICLE), emphasizes the need to consider the students' L1 as one factor affecting their writing (Granger, 1998). Thus, the proficiency level and L1 are important considerations.

Native speaker characteristics

In order to make principled comparisons, researchers have tried to ensure that there is as little variation as possible between the native speaker and learner corpus. In some cases this has meant choosing to use NS student writers rather than professional writers. Granger has headed the creation of a corpus of NS student writing with writings similar to those in the ICLE (1998) to use in comparison studies, and corpora used in other studies (Hyland & Milton, 1997; Reid, 1992) have

also followed this route. Other studies have compared the learner student corpus with a NS corpus written by professional writers (Flowerdew, 1998; Milton & Tsang, 1993). The appropriate choice depends on the research question under investigation.

Characteristics of the writing

Several characteristics of the writing--topic, the situation in which it was written, and the types of writing--all affect the features of the writing. In their creation of corpora, researchers have emphasized different qualities of the writing. The ICLE and its NS corpus equivalent both use argumentative essays, although the topics vary between the two corpora (Granger, 1998). Both the learner and NS corpora in Hyland & Milton's (1997) and Reid's (1992) studies were made of writings done as part of standard examinations. Flowerdew's (1998) study of causality markers used writing by students within a topic and register (essays about environmental concerns in the neighboring regions of southern China) which she felt would be more likely to contain features she was interested in studying. Her NS corpus was of writing in a similar topic, professional scientific writing in an environmental report. Milton & Tsang's (1993) study demonstrated some of the problems in finding comparable corpora. Because no one NS corpus matched their learner corpus with respect to register, the researchers pooled three different NS corpora. They justified this decision by saying that it gave a broader sample of NS writing and would thus help ensure that any differences between the learner and NS corpora indicated characteristics of the students' writing. However, they acknowledged that this corpus choice did have some limitations because the NS corpus was not taken from a

school context similar to the learner corpus and because the topics and writing genres were not similar. In my study, I will have to grapple with these questions of equivalent topics and registers.

Types of analysis

The majority of corpus linguistic studies of interlanguage are concerned primarily with underuse and overuse of certain forms by learners of English, as is my own study. According to Conrad (1999), however, corpus linguistics looks not only at the quantitative results of frequencies of use, but also takes a qualitative approach and makes interpretations about the functions of the forms and the significance of the contexts and situations in which they appear. Hyland and Milton (1997) demonstrated a clear example of combined quantitative and qualitative analysis in their study of qualification and certainty in student writing. They not only counted the frequency of occurrences of 75 words, but also randomly extracted sentence contexts of each lexical item from each level of student writing they were studying. They used these examples to determine the function and syntactical category of the words. An effective combination of both quantitative and qualitative analysis is an important consideration in my own research.

Summary

This chapter has introduced the modal auxiliary verbs I am interested in studying and given a framework of grammatical characteristics of form, semantic meanings, grammatical functions and relative frequencies of occurrence across registers. The modal verbs, especially in their epistemic or extrinsic functions, are crucial for academic and particularly scientific writing, where hedging, other politeness strategies, and the practical concerns of presenting an interpretation of a new finding with the appropriate degree of certainty make these forms, as well as other forms of modality, indispensable.

At the same time, modal verbs present a considerable challenge to learners of English, due to their unique grammatical characteristics, nuances of meaning and variance of use according to context and cultural interpretation. Past research has found modal verbs to be underused, overused and misused by learners in different situations.

Finally, my chosen method of study, corpus linguistics, has been discussed. Past researchers have grappled with corpora choices and degrees of analysis. General standards of corpus creation, finding one large enough and varied in a principled manner, are important considerations, as is the combination of qualitative and quantitative analysis. My application of these principles will be the focus of the next chapter, Methods.

CHAPTER 3. METHODS

In this chapter the creation of the two corpora used in this study, one learner language corpus and one professional corpus, is discussed. The characteristics of the student writers, their English proficiency, first languages, and areas of study are described, as are the types of writing and writing topics which make up the two corpora. I also list the modal verbs counted and describe the quantitative analysis, as well as the method of qualitative analysis and comparison made of the verbs occurring in the corpora.

The Learner Corpus

Two corpora were created in order to study learner auxiliary modal use in scientific research articles: a learner corpus and a professional corpus. The learner corpus was made of writings by non-native speakers of English who were international graduate students at Iowa State University. The 29 participants were all English 101D students in the 1999 calendar year. All subjects gave me their consent to be involved in the study, after I had obtained permission from the Human Subjects Committee at Iowa State University.

These students were admitted as regular full-time students at Iowa State and were enrolled in courses for their respective program of study. They were required to take English 101D because of their scores on an English writing placement test taken at the beginning of their studies at Iowa State. English 101D is an English

writing course for non-native speakers who are graduate students. The course focuses on scientific research writing, with emphasis on the organizational structures and typical grammatical features common in the research writing which graduate students will have to do during their course of study. Writing assignments include summaries and critiques of professional research articles, literature reviews, proposals and small research articles.

The students were a roughly even mixture of male and female who were mostly in their 20s. Their time in the United States or any other English speaking country ranged from six years to under two months. They came from a variety of language backgrounds, but the majority, 19, were Chinese. Five were Korean, three Turkish, one Brazilian, with a native language of Portuguese, and one Japanese.

These students represent the variety of international graduate students at Iowa State as well as at many other American universities. Their uniform level of ability, as determined by their placement into English 101D, ensures greater consistency in the results. Their diversity in background ensures that variation due to first language and culture is included in the corpus.

Characteristics of the writing

The writings which made up my learner corpus were all research articles written by the subjects as the final assignment for the English 101D course. Because these students hope to publish their research findings in respected academic journals, a genre exemplified by the published articles which make up the professional corpus, the research article assignment is an appropriate focus for this

study. All the papers were research articles in the sense that they presented original empirical data discovered at least in part by the writers, some also included data from colleagues in research group.

The student papers were all similar in rhetorical structure, with an Introduction containing background knowledge from past research, and a Methods section. In many papers the Results and Discussion sections were combined, leading me to combine the two in my corpus. The papers varied in length from 385 words to 3234 words. The total word count of the learner corpus was 46,728.

About half of the papers were first drafts, with no prior comments by instructors, and the other half were final drafts, which had been revised to varying degrees based on instructor comments. I collected some of these papers electronically, via disk and email, but the majority I scanned from the hard copy. After scanning, I edited the texts for any scanning errors, and then I divided each article into sections: Introduction, Methods and Results/Discussion.

The students' academic disciplines, and hence the subject of their papers, varied. Twelve wrote about topics in the natural sciences: five in genetics, one in veterinary medicine, two in chemistry and two in physics and two in food science and human nutrition. Several (10) also wrote in engineering and technology fields, four in aerospace engineering, three in material sciences and engineering, one in electrical engineering, one in computer science, and one in industrial technology. Four of the subjects were from the disciplines of consumer science and education: three were in textiles and clothing and one in nutrition education. Finally, two of the papers were within the field of economics and one was in statistics. (See Table 3.1

for a summary of the corpus characteristics.)

This wide range of subjects allowed my corpus to be more representative both of the range of backgrounds which international students come from, and the varied disciplines in which they hope to achieve accepted professional status, through, among other routes, publication of their own research. Using this wide and representative range should help to make my findings both more generalizable and more applicable to the a variety of students and TESL teachers.

The Professional Corpus

The writers

As I mentioned in the literature review, other researchers of NNS language have been divided about the choice of a comparable NS corpora. Some researchers choose to use writings by student writers, while others choose expert writers' work. I chose expert writers because I believe their writing serves as an important model for international students, scholars and scientists who want to enter this professional discourse community. That is, NNS students are not trying to write like NS students, but like published professionals. Some of the professional writers were non-native speakers of English, as are many scientists, but all the articles were edited to comply with the standards and expectations of the readerships of the British and American scientific journals in which they were published.

The texts

All of the writings which made up the professional corpus were already electronically available as parts of several different corpora: The Longman Spoken and Written English Corpus, (see Chapter 1 of Biber et al, 1999) the ARCHER Corpus, (see Atkinson, Biber & Finegan, 1994) and The Corpus of Writing in the Disciplines (see Conrad, 1996). In an attempt to match the genres of both corpora as closely as possible, all the writings which I chose were within the genre of research articles, presenting original findings of empirical data, with a rhetorical structure which could be divided into at least three sections: Introduction, Methods, Results and Discussion (often together). The total word number of the professional corpus was 419,864 words.

I wanted to match the topics of the research articles as closely as possible with the topics in the learner corpus; however, this was one of the most difficult aspects of the creating valid and principled corpora. A professional corpus with the same proportion of equivalent topics was impossible to create. Many of the students wrote about topics in engineering and computers, but there were few similar topics in readily available electronic corpora of professional texts. Many of the articles in the professional corpus had been published in medical journals, or were focused on other biological topics. Thus, the professional corpus was more weighted to topics from biology and medicine, while the student corpus was more weighted to areas of physics and engineering. I decided that similarity in genre was more important than similarity in topic, because of the relationship of research article writing with use of modality. Thus, these two corpora represent a variety of writers and topics, but a

Table 3.1. Description of the Student and Professional Corpora

Corpus	Number of texts	Number of words	Topic areas
Student	29	46,728	Natural Sciences; Engineering and technology; Family and Consumer Sciences; Social Sciences
Professional	134	419,864	Medicine; Natural Sciences

consistency in genre. However, this difference in topics was a factor to consider in interpreting the findings. (See Table 3.1 for a summary of the professional corpus.)

Quantitative Analysis

Both corpora were grammatically tagged for part-of-speech information, using the tagger developed by Biber (1988). An automatic program was then used to count all occurrences of the following items, identified as modals auxiliaries, marginal auxiliaries and semi modals, by grammar studies: *can, cannot, could, going to, has to* (and other tense forms), *may, might, must, needed to* (and other tense forms), *shall, should, supposed to, will, and would*. *Used to* is also considered a semi-modal; however, I found that very few of the instances for this phrase in my corpora functioned as modality, so I deleted this semi-modal from the study.

Counts were made by computer of individual as well as total modal use. Counts were also made by computer of total modals and individual modals in each research article section. The counts of each file were checked for any outliers by me. Files of the modals in context were also checked by me to assure that the modals had been correctly identified.

The counts were normed to 1000 by computer to make comparison across

texts of different lengths. I compared these counts of the different modals in both corpora and across research article sections in order to see the general frequencies of use and differences in these frequencies. Specific modal verbs with the greatest differences in counts--*can*, *may* and *will*--were identified for further analysis.

Qualitative Analysis

I looked at the surrounding 40-word context of all the occurrences of *can*, *may* and *will* in the student corpus and 100 randomly chosen occurrences of each modal in the professional corpus, similar to Hyland and Milton's study (1997). To look at the context, a computer program called Key Words In Context extracts a key word, and the 20 words before and the 20 words after this key word. In my functional analysis, I loosely following Butler's (1990) categories of analysis, which I chose because Butler's study was the only one to look in depth at modal verb use and to do semantic and functional analysis of the forms. I identified *can* and *may* as falling into one of the following semantic meanings: epistemic comment, legitimacy, ability, ambiguous meaning, or another meaning unidentified by Butler or other researchers. Uses of *will* were categorized as meaning one of following: future time prediction, volition, induction or deduction. Counts of occurrences of each of these different semantic functions were made by hand, and differences in counts between the two corpora were noted. Where I saw important differences in the prevalent functions of each verb, further qualitative analysis was done to determine possible reasons for the differences. Finally, I tried to make connections between the different proportional functions for which modal verbs were being used and the overall

quantitative findings. Combining quantitative and qualitative findings provides a picture of frequencies of student use and reasons for the differences in frequencies.

Summary

In order to answer the research questions regarding how students used modal verbs differently from professional writers in scientific research articles, I created two corpora, a student corpus of research articles from a variety of disciplines, written by international students in the English 101D course at Iowa State University, and a professional corpus of research articles in different fields taken from a variety of previously created corpora. Modal verbs were counted in each of the corpora, and the occurrences of the modal verbs in both of the corpora and in each of their research article section divisions were compared. The modal verbs with the greatest differences in counts--*may*, *can* and *will*--were analyzed further to find differences in their contextual use and semantic functions within the two corpora. These findings are presented and discussed in the following chapter.

CHAPTER 4. RESULTS AND DISCUSSION

This study examined the differences in the use of modal verbs in scientific and research articles by NS professionals and advanced international graduate students. It looked at this difference within the context of the overall frequency of modal verbs used in research writing, as well as the frequency differences between research article sections and between individual modals. A further goal of the study was to examine the functions of modal verbs used by the two groups in order to note any similarities or differences.

In order to address these questions, this chapter has two basic sections. The first section focuses on the differences in modal verb counts for the two groups, including differences between research article sections and in the use of a few of the more important individual modal auxiliaries. The second section discusses the possible reasons for differences in counts of modal verb use between students and professionals, including the different functions modal verbs serve for each group of writers.

Quantitative Findings

Overall counts

This section presents the total number of modal verbs used by each group. As described in Chapter 2, Literature Review, the modal verbs counted were *can*, *could*, *may*, *might*, *must*, *shall*, *should*, *will*, *would*, and the semi-modals *going to* and *has*.

to, *needed to*, and *supposed to*. This list was developed on the basis of current understanding of modal auxiliaries (Biber et al, 1999).

The counts immediately show obvious differences between the means of the two groups. While the professional corpus had a mean count of about 6 modal auxiliaries used per 1,000 words, the student corpus use was almost twice as frequent, with a mean of 11.95 modal verbs used per 1,000 words.

The professional count is somewhat lower than past corpus studies of modal verb use (see Table 4.1). However, the difference may be related to differences in the corpora. Biber et al's (1999) corpus of academic writing was not limited to scientific research articles, and Butler's (1990) corpus was made of not only research articles but also college-level textbooks. A comparison with studies of corpora closer to my own, which looked at only at counts of specific modal verbs, suggests these findings are not overly low (see Table 4.2). In Varttala's (1998) study of medical research articles and articles from more popular scientific journals, occurrences of *may*, *could*, and *might* were counted, and 4.16 occurrences per 1,000 words were found, compared with a count of 3.0 of these verbs in my corpus. Hyland's (1996) study of research articles reported 3.58 occurrences per 1,000 words of *could*, *may*, *might*, *should* and *would*, compared with a count of 4.0 of these verbs in my professional corpus. In the corpora which were the focus of the present study, we see the professionals use about six modal verbs per 1,000 words, and students use twice as many modal verbs as the professionals do.

Table 4.1. Studies of overall modal verb use in professional corpora

Studies	Number of modal verb occurrences per 1,000 words
My study of modal verbs in research articles.	6
Biber et al (1999) corpus of academic writing	~13
Butler's (1990) corpus of journal articles and college textbooks	9.56

Table 4.2. Studies of use of specific modal verbs in professional corpora

Studies of modals	Counts per 1,000 words	Counts of the same modals in the present study
Varttala (1998) <i>may, could, might</i>	4.16	3.0
Hyland (1996) <i>could, may, might, should, would</i>	3.58	4.0

Range in overall counts

Besides the obvious difference in frequencies, the student corpus exhibited a greater range in the number of modals used from article to article than professionals did, as illustrated in Table 4.3. The high count in the professional corpus was 17.7 modal verbs per 1,000 words and the low was zero modal verbs per 1,000 words. However, the range in the student corpus is more striking. The high number of modal verbs used in these 29 papers is 25.26 per 1,000 words in one paper, while the paper at the low end of the spectrum used 0.76 modals per 1,000 words. Both corpora contain research articles about a variety of subjects, which may contribute to some of the differences in modal use between individual articles. Another reason for the differences may be related to the fact that half of the papers were first drafts and the other half were revised based on instructor's comments. Related differences in the teaching styles and emphases may also be a factor, as may be the relative skill

levels of different students.

Table 4.3. Means and Ranges of Modal Verb Use

Corpus	Mean counts of modal	Range	Number of papers/files	Stan. Dev.
Professional	6.5	0-17.7	136	3.88
Student	11.98	0.76-25.26	29	6.6

Research article section counts

This section breaks down the number of modal auxiliaries used by research article section. Results from both corpora concurred with the past research which suggests Methods sections have the lowest occurrences of modal verbs and Introductions and Discussion sections have the highest counts (Adams Smith, 1984; Butler, 1990; Salager-Meyer, 1994; Skelton, 1997; Varttala, 1998). Previous research has looked at Results and Discussion sections separately, but because, as noted in Chapter 3, students often did not break their treatment of results into separate Results and Discussion sections, I was not able to notice differences in the counts of these different sections. In Methods sections, both students and professionals used fewer modals than in their Introductions or Results/Discussion sections. Students used more modal verbs in Results/Discussion sections than in Introduction sections (Figure 4.1).

Despite the similarity in the proportional use of modal verbs in research article sections, students demonstrated a consistently higher frequency of modal use throughout the sections. The difference between the two was most pronounced in the Methods sections. In Methods sections, students used almost four times as

many modals as the professionals did (6.5 occurrences per 1,000 words compared with 1.7 occurrences per 1,000 words). Students again used almost twice as many modals as the professionals in the Results/Discussion sections (14.7 occurrences compared with 7.4). In the professional corpus, modal verbs occurred with the same frequency in the Introductions and Results/Discussion sections. However, students used more modals in the Results/Discussion sections than in the Introduction sections. In the Introductions, students used 1.5 times as many modal verbs as the professionals did (11.9 occurrences compared 7.4). Figure 4.1 illustrates the total modal counts in each research article section. We can see that students use more modal verbs than their professional counterparts throughout the research article sections, with the most pronounced difference occurring in the Methods sections.

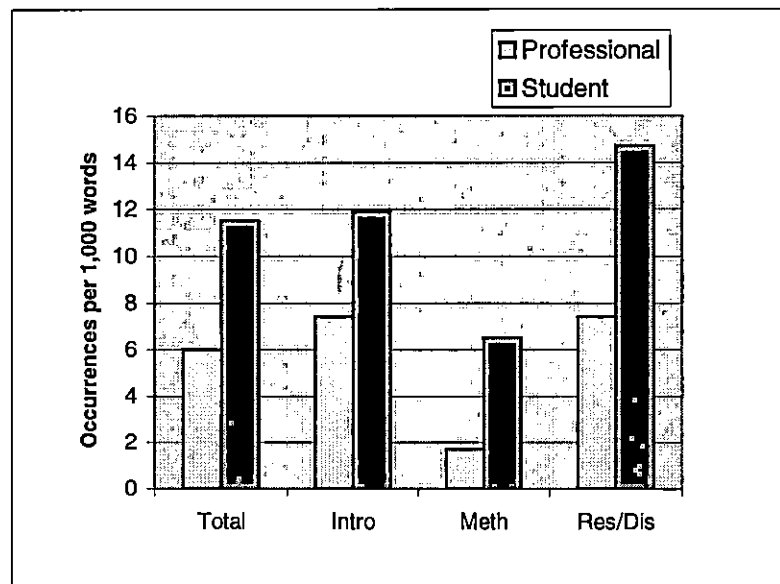


Figure 4.1 Occurrences in research article sections

Ranges in section counts

The range of occurrences in individual sections of the student corpora is more extreme than the range in the professional corpus, as illustrated in Table 4.4. In the Introductions, the range of modal occurrences among individual papers in the professional corpus went from a low of 0 to a high of 26.8. In the student corpus the equivalent range was 0 to 42.4 occurrences per 1,000 words. In the Methods sections, professional papers differed within a range of 0 to 9.3 occurrences, while students had a range of 0 to 26.8 occurrences. And in the Results/Discussion sections, students also exhibited a greater range, from 0 to 43.3, compared with the professional range of 0 to 29.4. In each corpus and in each section, the low frequency of occurrence was always zero, but the students typically demonstrated more extremes in range.

Again, several factors may influence the greater range in student counts, including the variety of subjects, differences between first and revised drafts, different instructors and different skill levels.

Table 4.4. Ranges of Modal Counts in Article Sections

Corpus	Intro	Meth	Re/Dis
Professional	0-26.8	0-9.3	0-29.4
Student	0-42.4	0-26.8	0-43.3

Counts of individual modals

This section focuses on the differences in counts of three modal verbs that showed notable variation between the two groups. While there were small differences in the frequencies of occurrence in almost all of the modal auxiliaries, there were only a few instances where the differences were greater than 0.5 occurrences per 1000 words. The modals *may*, *can*, and *will* showed the greatest difference throughout the corpora and thus account for a large part of the difference in counts between the two corpora, although *should* and *could* had a differences greater than 0.5 in Results/Discussion sections. *May* was the one modal verb which was consistently underused by the learners, while *can* and *will* were examples of some of the greatest overuse of modals, and looking at these examples may give us a greater understanding of the causes for differences in use between the two corpora.

Occurrences of *may*

May was used less often by students than by the professionals. As Figure 4.2 shows, the student corpus had half as many occurrences of *may* as the professional corpus did, with 1.7 fewer occurrences in the Introductions, or about one third as many, and half as many in the Methods sections. The difference in the Results/Discussion sections was not as pronounced. Students used 80 percent of the *mays* that professionals did, 1.9 occurrences compared to 2.4 occurrences.

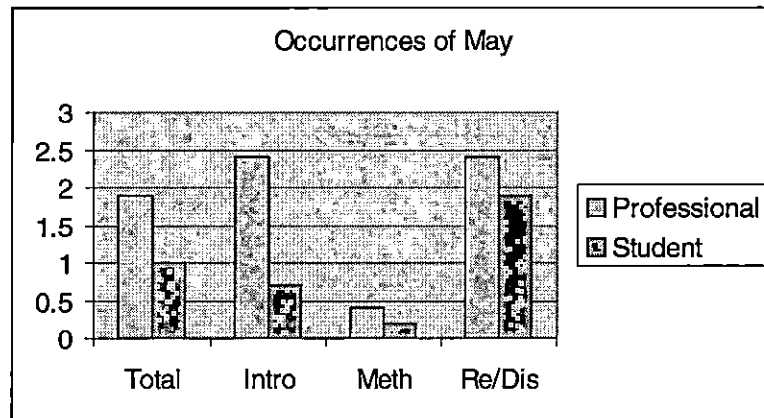


Figure 4.2. Occurrences of *may* per 1,000 words in each section

Occurrences of *can*

As illustrated in Figure 4.3, the professionals used 1.3 *cans* per 1,000 words in the total corpus and students used 4.7 *cans* per 1000 words, 3.4 more occurrences per 1,000 words or 3.6 times as many occurrences by the students. In the Introduction section, professionals used 2 *cans* per 1,000 words and students used 4.2 *cans* per 1,000 words, 2.2 more per 1,000 words or about twice as many. The differences in use of *can* in the Methods sections were most striking. Students used 19.5 times as many *cans*, 3.9 occurrences compared with 0.4 occurrences in the professional corpus. In the Results/Discussion section professionals used about 1.5 occurrences of *can* per 1,000 words, and students used 5.6 occurrences--3.7 times as many.

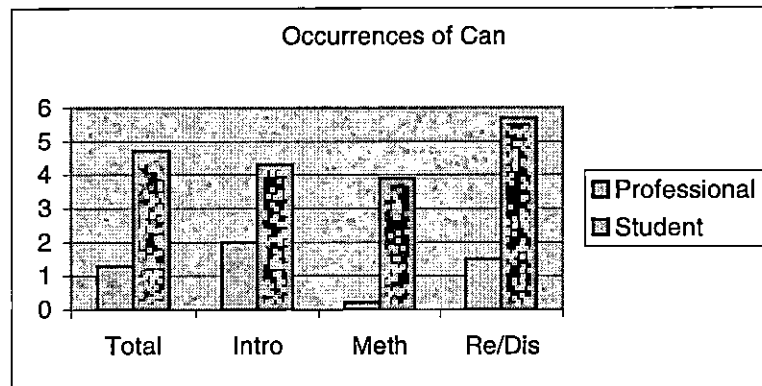


Figure 4.3. Occurrences of *can* per 1,000 words in each section

Occurrences of *will*

As illustrated in Figure 4.4, *will* had 0.9 more occurrences in the total student corpus, or 5.75 times as many occurrences in the student corpus than in the professional corpus. In the Introduction section *will* was the modal which was most overused. While professionals used only 0.3 occurrences of *will* per 1,000 words in the introductions, students used 3.6 occurrences, twelve times as many. There was also a noticeable difference in the Methods section, where students used 14 times more occurrences of *will* than professionals, (1.4 compared with 0.1). Students also used more than twice as many *will*s in the Results/Discussion sections, where the student corpus had 1.6 occurrences per 1,000 words compared with the professionals' 0.6 occurrences.

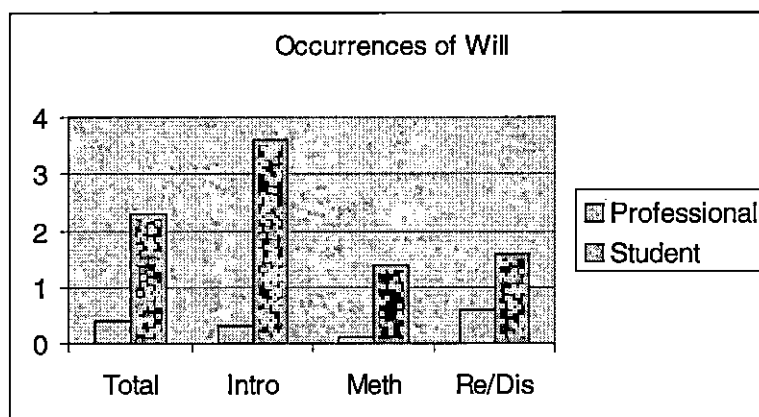


Figure 4.4. Occurrences of *will* per 1,000 words in each section

We see from this section that differences in modal verb use between the two corpora were most significant in the instances of three different modal verbs: *may*, *can*, and *will*. *Can* shows the most prominent differences, while *may* shows the only evidence of underuse of any modal by the students. Two other modals: *could* and *should* also had differences greater than 0.5 in the Results and Discussion sections. The student Results/Discussion corpus had 0.9 more occurrences of *could* and 1.0 more occurrences of *should*. The uses of these verbs are not covered within the scope this study, but remain important areas for further research.

Functions of the Modal Auxiliaries in the Corpus

An overview of common semantic categories of *may* and *can*

The previous studies reviewed in Chapter Two showed that *may* and *can* have potentially similar meanings and functions. Both can be used extrinsically to

make an epistemic comment about the truth, probability or possibility of the truth of any statement. *Can* and *may* may also be used intrinsically to make a statement about the ability of a certain agent. The results suggested that *may* was more likely to be extrinsic than intrinsic; however, all the possible functions for both *can* and *may* were found in my corpus, although in varying degrees. In this section, I exemplify these uses before going on to discuss the differences in use between the two corpora for each of the modals separately.

In this example of a common use of *may*, the modal verb serves an extrinsic, epistemic function in the professional corpus:

Because most height growth occurs early in stand development, sapwood volume and maintenance respiration *may* not increase greatly between intermediate and old-growth stands.

May can be paraphrased as "It's possible that sapwood volume and maintenance respiration do not increase greatly...." Without the modal meaning the proposition is expressed with no qualification and conveys a more direct, positive meaning:

Because most height growth occurs early in stand development, sapwood volume and maintenance respiration do not increase greatly between intermediate and old-growth stands.

Can serves the same function, expressing possibility, in this example from the professional corpus:

Screening of surface charges by monovalent ions *can* be significant, particularly when divalent ion concentration is low.

Unlike the above extrinsic meanings, the intrinsic meaning of *can* and *may* imply ability or permission. In the following example of *may* in an intrinsic function, from the student corpus, the sentence can be paraphrased, "The whole series...is often able to be synthesized with the same crystal structure....":

The whole series of intermetallic compounds r-t-sb *may* often be synthesized with the same crystal structure, differing only in rare earth elements.

Can much more commonly serves this intrinsic function, as in this example in the professional corpus:

Despite these characteristics and the observation that myc can modulate gene expression, a direct role for myc or Max as transcription factors has never been demonstrated.

Can and *may* can also be used in an intrinsic manner to serve a more extrinsic function. In these cases *may* and *can* are used to suggest legitimacy of a viewpoint or logical possibility, as in these examples from the professional corpus:

The earlier paper *may* be regarded as essentially the introduction to this paper and here we just note one or two points about the....

The tendency for ex ratios of adults to be skewed toward females, therefore, *can* be attributed to a differential survival.
In academic writing, the distinction between extrinsic and intrinsic functions

(meanings of possibility or ability) is not always easily determined. The meaning is often ambiguous, able to serve both functions, sometimes logically able to mean both ability and possibility at the same time, and in other occasions, able to mean

only ability or possibility, although the intended meaning is unclear (Coates, 1983).

In this example from the professional corpus, *may* could serve to indicate both that bile acids are able to damage DNA and that it is possible that bile acids damage DNA:

"Bile acids *may* directly damage DNA, or chromatin."

A similarly ambiguous meaning of *can* is seen in this sentence from the professional corpus.

We conclude that the mechanism will be effective in low-viscosity basic and intermediate magmas and *can* occur in basaltic magmas for crystals of diameter 1mm.

This sentence could be rewritten as either: "...and is able to occur in basaltic magmas," or: "...and possibly occurs in basaltic magmas."

I chose not to classify these ambiguous meanings as either extrinsic or intrinsic, rather, I placed them in their own "ambiguous" category.

Throughout both corpora, we see the modals *can* and *may* are used for their traditionally accepted functions, to indicate possibility, ability, or legitimacy, or a more ambiguous meaning. The differences between the two corpora are differences of degree and proportion rather than a radically different understanding of the use of one or more verb.

Epistemic use of *may* most common

Although the student and professional corpora exhibited similar functions of *can* and *may*, a look at the proportional use of these functions reveals some of the

differences between the corpora. In both the professional and student corpora, the use of *may* occurred almost solely in an extrinsic, epistemic function, as in this example from the student corpus:

This *may* provide a direct method for making an integrated, planar, focusing transducer using existing integrated circuit fabrication technology.

Because *may* was so overwhelmingly used extrinsically, this function is the only function of *may* discussed in depth in this study. The findings suggest students who use *may* understand the common function and semantic meaning associated with the word. However, as noted in the quantitative findings, students use *may* considerably less often than the professionals, especially in the Introduction sections, where the student corpus had 1.5 less occurrences per 1,000 words than the professional corpus. The biggest differences between use of *may* in the two corpora was in the Introduction sections of the corpora. In Introductions, extrinsic statements of certainty and doubt are less common when reporting about others' research (Thompson and Yiyun, 1991) but occurrences of *may* are seen in the introduction of a controversial area within the field or a phenomena that still has plenty of questions. In other words, *may* and other forms of expressing uncertainty are used to discuss the areas of necessary additional research or the "niche" in the research (Swales, 1990) that the researcher wants to expose and to address.

In these Introductions we see professionals discussing some unknowns in their field through the presentation of what *may* be or possibly is the case:

Phytomass incorporates physiologic controls over growth as well as

evolutionary responses to field conditions, making it difficult to sort out proximal and ultimate controls. For example, growth of arctic and alpine plants *may* halt while there are still resources available, which is presumably an adaptation to unpredictable growing season length.

Selective pressures from natural enemies *may* also determine host use. Certain hosts *may* provide "enemy-free space." Some hosts *may* provide substrate where enemies are absent or in low abundance. Other hosts *may* provide protection by furnishing a background for crypsis or by supplying compounds acquired by the herbivore that are not tolerated by natural enemies. Thus plant chemistry *may* be important in mediating the interactions between herbivores and their natural enemies. We sampled for the absence of enemies or the reduced impact of enemies with the use of certain hosts.

Unlike the professional corpus, the student papers in the study often did not make many tentative statements in their Introductions that would indicate the room for questions or doubt in their field. The students were more likely to merely report what past research had indicated, as this example from a paper in genetics indicates:

Using another isoenzyme, peroxidase, Yang and his colleagues successfully test corn hybrid seeds purity. But isoenzymes are easily affected by environmental factors, In many cases, the isoenzymes-based purity tests are unstable and irreproducible.

Advances in molecular genetics allow to develop more effective and stable DNA-based seed quality assessment approaches. Among these methods, RAPD is the most popular. Chen applied this technology to check the rice hybrid-xianyou 63 seeds purity. The result was the same as that of seed germination. Luan also used RAPD to test tomato hybrid seeds purity successfully.

Although RAPD overcome the drawbacks of isoenzyme-based approach, it is still time-consuming, and costly. In this paper, we report a new DNA marker method, which is more suitable for large-scale seed purity tests.

This learner does not use modal verbs or any other hedging devices to the to explain the complexities of the research question.

This underuse of *may* suggests that throughout their writing, students write overly positive, confident statements of fact, when a statement which is less confident, suggesting possible rather than positive truth of a proposition, would be more appropriate. For example, in the following selection from the Discussion section of an economics article about the effects of welfare reform, the student writer makes almost no qualifications to the strong statements made in assessing the situation:

Therefore, the success of welfare can be accomplished with enough low-skilled jobs and proper wage that can provide resource for needs. People who stay in welfare are less educated than people who leave public assistance, so another strategy of success is providing job training and higher education. Because children under age three make difficult for a parent to work, proper child care can be an incentive to work.

The statements in this article could be interpreted as too definite to be credible. Even the use of *can* in the last sentence is more suggestive of positive ability rather than a more qualified statement of a possible effective strategy. Within the entire Discussion section, no instances of *may* occur. Consider the different, possibly more credible attitude conveyed when *may* is used to suggest possibility:

Therefore, the success of welfare *may* be accomplished with enough low-skilled jobs and proper wage that can provide resource for needs. People who stay in welfare are less educated than people who leave public assistance, so another strategy of success *may be* providing job training and higher education.

In this version, the student's proposition is conveyed as likely but not definite, a stance which is easier for a reader to accept than the first, unqualified statement.

Thus we see students using *may* for the same function but to a lesser

degree than their professional counterparts, resulting in writing which conveys a more direct, less qualified attitude, and neglects to express the unknowns in the writer's area of research or the uncertainties inherent in almost every claim.

Semantic classifications of *can*

This section introduces the various semantic classifications of *can*. Like *may*, *can* functions in both extrinsic and intrinsic forms. Extrinsically it serves to comment on the possibility of a proposition. Intrinsically its possible meanings include ability and legitimacy or permission. *Can* at times becomes much more difficult to analyze after these more obvious meanings have been established. As Biber et al acknowledge, "*Can* is especially ambiguous in academic prose, since it can often be interpreted as marking logical possibility or ability" (1999:492).

Most of the professional occurrences of *can* were relatively simple to classify as either speaking about legitimate logical possibility or about ability. This example from the professional corpus demonstrates the use of *can* to indicate ability:

For example, translocation of fixed carbon away from source tissue *can* reduce feedback inhibition of photosynthesis resulting from photosynthate accumulation under high carbon dioxide.

Another professional use of *can* serves to communicate logical possibility, an epistemic function:

However, some recent field experimental evidence suggests that biotic interactions also *can* be important to grasshoppers.

A third use of *can* by professionals indicates legitimacy of a certain way of viewing or understanding something--a type of permission:

The three properties *can* be explained on the assumption that a synapse will be potentiated...

Ambiguous uses of *can* were rare in the professional corpus, but occurred often in the student corpus. Ambiguous meaning is exhibited in this example from the student corpus, in which the student's meaning could be interpreted as logical possibility or ability:

"An new vector *can* be computed as follows."

Can serves a range of functions, some of them more easy to interpret and classify than others, in both the student and professional corpora. All of these functions were found in both corpora. The different ways in which *can* was used in the student corpus, to express possibility, ability, legitimacy, or a more ambiguous meaning, are discussed more extensively with examples below.

Student use of *can*

May is the only modal that students consistently underused, while their use of *can* was the most obvious example of overuse of modals. This section discusses the functions *can* served in student overuse of the verb. Students used 3.3 more occurrences of *can* per 1,000 words in the total corpus than the professional writers did, and students used 3.7 more occurrences of *can* per 1,000 words in the total

corpus than they used *may*. The students' overuse of *can* is immediately obvious in their writing, as it is in this example from a student of consumer science.

Many consumers have thought if they *can* access the Internet easily, they'll often use the Internet as a medium to search or purchase. It implies that if they *can* use the Internet easily, such as at home, workplace, and public library, they'll use it more than if they *can* not access it. That is difficult access Internet *can* be one of the barriers to using Internet shopping.

Can* substituted for *may

To some degree the overuse of *can* may help to explain the underuse of *may*. *Can*, rather than *may*, often functions in the student corpus to propose the areas of uncertainty, as this example from an introduction by a student in veterinary medicine indicates:

Several works suggest that glial cells *can* be important source of EAAs. Szwatkowski (1990) proved that the Muller cells, type of retinal glial cells *can* release glutamate through reversion of glutamate transporter activity.... Since bradykinin is the main product of the damage of blood vessels we have asked whether this agent *can* cause glutamate release in the retinal glial cells as a response on the damage of the retinal blood vessels in glaucoma.

As we saw earlier in the professional corpus, Introductions often contain *may* to express areas of uncertainty; the use of *can* in such a role distinguishes this student writer's introduction of the niche from the professional writer's Introduction.

***Can* more often intrinsic**

Although in the above example the student's use of *can* functions to indicate possibility, *can* is still used intrinsically. Each of the instances of *can* in the above sentences can be restated with "are able to be," although the first *can*, "glial cells

can be important source..." has a more ambiguous semantic function. Both the professional and the student corpus have far more instances of *can* with meanings of ability and legitimacy than of epistemic possibility. This suggests that *can* is regarded as generally serving a different function than *may* by both groups of writers.

***Can* used to express legitimacy**

While in the professional corpus the most common use of *can* was to suggest meanings of ability, and secondly meanings of legitimacy, in the student corpus just the opposite was true. Meanings of legitimacy, or meanings that suggested both legitimacy and ability were more common than simple meanings of ability. Students often used *can* to suggest the legitimate or logical way to view a proposition, as these examples demonstrate:

Therefore, these types of research *can* be viewed a stepping stones to link nutrition education with desirable food-and-nutrition-related purchasing behaviors.

We *can* see the oxidation rate of all compositions were pretty low.

Through their use of *can* as an expression of legitimacy, they seem to have substitute the more common (to the NS) way of expressing doubt or certainty about a proposition, through the use of *may*, for an expression which made the reader and the writer a more immediate part of the framing of the truth of the proposition. This use of *can* either directly, in the active voice, or indirectly, in the passive voice, suggests a human agent, something which the use of *may* in a extrinsic, epistemic

function does not do.

The statement:

From the above examples, we *can* say that the modification and the implementation can work well.(from the student corpus)

indicates a human agent "we" and thus puts more of the responsibility for assessing the truth of the situation directly on the writer and reader than the statement:

For example, growth of article and alpine plants *may* halt while there are still resources available. (from the professional corpus).

Consider the student statement rewritten with *may* rather than *can* in the legitimate function:

Modification and implementation *may* work well, based on the above examples.

In this rewrite, the presence of the writer has been removed, and the more distanced writing suggests the objective scientific writing we are used to.

Throughout the corpus we see students frequently using *can* intrinsically to convey a permissible or legitimate action performed by an often specified agent, such as "we *can* say," or "that *can* be attributed."

Ambiguous use of *can*

Besides overuse of *can* to indicate legitimacy, learners also often overused *can* for purposes of ambiguous meaning. While ambiguous meanings of *can* in academic writing are to be expected to some degree (Biber et al., 1999) the uses of *can* for meanings which could not easily be classified was by far a much more common feature of the learner corpus than the professional corpus. These uses of *can*, which could be interpreted as meaning legitimacy and ability and at times even possibility, often occurred in physics and engineering texts, where the discussion of mathematical formulas seemed to encourage a use of *can* that fell somewhere between the two meanings. The following examples from the student corpus demonstrate this ambiguity between legitimacy and ability:

So we *can* caliber the darkness reader within each part of the curve
and find the thickness of the material by linear interpolation.

At the same flow angle α , we *can* see when flow was inviscid, forces were zero, when flow was viscid, forces were not zero.

These could be interpreted as meaning either, "So we are able to caliber the darkness reader," or "So we legitimately may caliber the darkness reader...."

Again we see *can* being used with a human agent, "we *can* ...," creating a tone not common in the professional corpus. The following examples show more ambiguous uses of *can* which suggest a use not found in professional writing.

All electrons are negatively charged, so it might seem surprising that two like charges *can* attract.

Based on work , it is feasible that integrated pvdf fresnel array

transducer *can* be realized.

If the unemployment rate of country where recipients live is lower, they *can* have higher chance to get a job and to get higher wage.

One of the opamp inputs must be connected to one of the bipolar transistors so that a Ptat voltage or current *can* be produced.

The use of *can* in these sentences causes them to seem marked and unnatural. In professional writing *will* or *may* or no modal auxiliary at all, depending on the degree of tentativeness, would be more likely to appear.

Consider the above examples rewritten to use *will* or *may* or written without a modal:

All electrons are negatively charged, so it might seem surprising that two like charges *will* attract.

Based on work, it is feasible that integrated pvdv fresnel array tansducer *will* be realized.

If the unemployment rate of country where recipients live is lower, they *may* have higher chance to get a job...

One of the opamp inputs must be connect to one of the bipolar transistors so that a Ptat voltage or current is produced.

The rewritten sentences seem less marked when *can* is removed.

This use of *can*, and more specifically use of *can* to speak about formulas, is most prevalent in students' Methods sections, especially the Methods sections of physics and engineering texts, as seen in these examples:

Equation 14 *can* be true if the measured modes t_j is the linear combination of the columns of the L_{ij} .

For a cmos process with $|v_{tp}| = 0.7$ v , the minimum supply voltage *can* be lower than 1 V with a value approximately equal to the theoretical limit of the proposed bandgap reference .

To ensure v_b equal to $v_{bn} - v_{tn6}$, the input stage of the *tia* *can* be realized based on the same structure as the one that produces *l3*.

A look at these examples helps explain the drastic overuse of *can* in Methods sections.

The tendency of this type of use in physics and engineering texts, especially in Methods sections, suggests some of the difference in the frequencies of *can* in the student corpus could be related to the topical differences between the two corpora. More than a third of the student writings are about the subjects of physics, engineering and computer science, and not only do the students writing in these subject areas have more frequent occurrences of *can* than students in other areas, but past research has also shown that at least in physics *can* is used more frequently than in the writings in biology (Butler, 1996). Students in engineering and physics are more likely to use *can* in formulas and in the ambiguous function suggesting both logical legitimacy and ability, and these topics are less represented in the professional corpus, in which most of the writings are from the natural sciences.

However, this explanation does not explain the marked, unnatural tone many of these ambiguous uses of *can* carry, and this explanation does not account for all of the occurrences of *can* in the student corpus, as many non-engineering texts also had frequent occurrences of *can*.

Throughout the student corpus we see *can* being used in ways uncommon in the professional corpus. *Can* often is used by students to suggest possibility, a function served more often by *may* in the professional corpus. The students also use

can to suggest legitimacy, a legitimate or permissible way of understanding a proposition. Often students use *can* in ambiguous meanings, suggesting probability, ability and legitimacy.

Use of *will*

The student corpus also had a greater proportion of occurrences of *will* than the professional corpus, especially in the Introduction section. This section introduces the semantic categories of *will* found in the corpora. *Will* usually, but not always, implies a future reality. Its meanings can be divided along extrinsic (functioning as prediction) and intrinsic (functioning as volition) lines, but I felt this was too simplistic for my purposes, as almost all academic writing can be expected to use *will* for prediction, rather than volition (Biber et al, 1999). Butler (1996) used a more detailed method for dividing extrinsically used *wills* (originally developed by Huddleston, 1971). In this system, *will* is categorized according to one intrinsic meaning, volition, and three extrinsic functions, future time, induction and deduction.

In its intrinsic meaning, *will* is used to suggest volition, as in this example from the professional corpus:

(From a Methods section) The classes roughly correspond to suppressed, intermediate, codominant and dominant crow classes and *will* be referred to as such.

Extrinsically *will* often indicates future time, as in this example:

(From a professional Discussion section) ...there is little doubt that this field *will* expand into one of the major applications of Mri in oncology.

A second extrinsic function of *will* indicates induction, suggesting a reality that

is a recognized truth, as in this example:

(From a Discussion section) Moreover, the fermentation of fibre *will* have effects on colonic ph.

Finally, *will* also may indicate deduction, demonstrating a logical relationship between two statements, usually "if.....then" sentences. This function is exemplified in the following selection from a professional introduction:

If woody tissue respiration varies with surface area, then respiration *will* increase with stand age.

The differences in student use of these different functions of *will*, to indicate volition, prediction, induction and deduction, are discussed below.

Student use of *will* in Introduction sections

Will indicating volition

The students were far more likely to use *will* as an act of volition, especially in the Introduction section, than the professional writers, in which volition occurred rarely, usually only in the Methods section.

In this example from the student corpus, *will* serves to indicate volition,:

In section 2, we *will* introduce the H & 6 algorithm and our modification scheme.

In the student corpus, the personal pronouns *I* and *we* often co-occurred with *will* in volitional meanings, something that only occurred twice in the professional corpus, both in Methods sections.

(From the student corpus, Introduction section): I *will* write a

program to simulate the pull cellular manufacturing process. I *will* focus on what the change of manufacturing capability *will* be. I *will* use this program to collect data about the quantity of products.

Professionals' use of volitional *will* was more likely to co-occur with passive forms, as in this example from a Methods section:

These classes roughly correspond to suppressed, intermediate, codominant and dominant crown classes and *will* be referred to as such.

Making predictions

Students often used *will* in Introduction sections to indicate prediction, in order to imply that the proceedings of an experiment or the writing of the paper itself would continue according to their intention, as in this example from an Introduction:

In this paper the primordial phase of fmGA *will* be applied to a helicopter rotor problem where "building blocks" for better solutions *will* be selected from several thousands of designs. Among the types of mGA, fmGA was chosen to be studied for the helicopter rotor design problem because of its relatively small size of initial the initial population called 'probabilistically complete initialization' thus the small number of total function evaluations. The results *will* be investigated to see if fmGA can effectively select promising building blocks through its primordial stages even started with probabilistic initialization scheme.

When professionals wrote an overview of their work in Introductions, they did not use *will*, rather they used present tense or past tense, with no modals, to inform readers what was done in the study or to give a preview of the paper. These two examples of a professional overview in the Introduction demonstrate this type of overview:

This paper continues an investigation of food plant related variation in palatability of the Florida queen butterfly, *Danaus gilippus berenice* (cramer).

Because similar mechanisms may exist in other alpine species, we consider only live above ground phytomass here, but we use both the current and previous year climate data in our analysis. We used path analysis to examine phytomass climate relationships.

The above examples contrast the ways students give overviews of their papers, with *will* in volition and prediction functions, with the way professionals write overviews, without modal forms.

Both groups employ the future use of *will* to discuss hypothetical predictions and to discuss the future ramifications of their work or to make other future predictions.

From the professional corpus:

We propose that mip-1 and other cytokines with glycosaminoglycan-binding sites *will* bind to and be presented by endothelial proteoglycans to trigger adhesion selectively not only of lymphocyte subsets....

From the student corpus:

This work was conducted to have an idea about the response that the powders, which *will* be produced throughout this study, *will* give to the ultrasonic disruption with respect to their sizes and agglomeration among each other.

In the first example, *will* is used to present a hypothesis, and the second *will* in the next example serves to present the research question being studied.

Students were more likely in Introduction sections to speak strongly about the future with regard to the predicted effects of their work or other scientific findings. As the following examples demonstrate:

But this question is really important in manufacturing factories. It *will* affect the manufacturing cost, the factories layout, the quality of products...

If the electricity could be sent through superconducting cable, not a kilowatt-second of energy would be lost, billions of dollars *will* be saved.

In this way, students argued the importance of their work; professionals, however, used similar sentences not in the Introductions but in Discussion sections. The following examples from professional Discussion sections use *will* to talk about the future effects of their findings:

We propose these simpler diagnostic criteria *will* lead to earlier diagnosis and more effective treatment of this group of potentially lethal disorders.

It is hoped that this *will* produce a pattern characteristic for each strain of organism.

Not only do the professionals use these type of statements in the Discussion sections, where they have already presented the research and findings in detail, but they also hedge these statements with phrases like "*We propose...*" and "*It is hoped,*" something students didn't do. The result is student writing which sounds overly positive and confident before any of the actual research has been presented.

Throughout the Introduction sections we see students overusing *will* for functions where the professional do not use *will*. Students used *will* volitionally, to give an overview of the study, and to discuss the predicted effects of their research.

Student use of *will* in Methods sections

Students writing in the areas of physics and engineering often used *will* in Methods sections for reasons similar to those influencing their use of *can*. *Can* and *will* were both used in Methods sections to talk about engineering problems or mathematical formulas. While *can* was used to indicate how to form or solve a certain problem or formula (eg: "So the location of the damage *can* be found by searching for a value that is smaller than others in the matrix."), *will* was used to indicate the expected results of a formula or situation. The following examples from engineering Methods section demonstrate this use of *will*:

By doing this, it *will* reduce the voltage difference between those nodes and hence, the effects of offset voltage.

If the measured vector already lies in the subspace, then Oakj and okj *will* be different.

The very few occurrences of *will* in Professional Methods sections are used to express the intentions of the researchers and occasionally to express an induction similar to what we see in the student use. The following examples from the professional corpus demonstrate professional use of *will* in Methods sections:

These classes roughly correspond to suppressed, intermediate, codominant and dominant crown classes and will be referred to as such.

The monoclonal antibody pc10 *will* "recognize" pcna in conventionally fixed and processed histological material.

The first example demonstrates use of *will* to express intention, and the second expresses induction.

Students consistently used *will* more in Methods sections than their professional counterparts, especially in engineering and physics papers, where *will* is used to express inductive logic in formulas.

Professional use of *will* in Discussion sections

Professionals also used *will* more often to discuss the future needs in their field, a function usually occurring in the Discussion section, and one uncommon in the student corpus. These examples of *will* in the professional corpus demonstrate the use of *will* to indicate future needs:

Further high-resolution observations *will* be needed to distinguish between these possibilities.

...y-linked marsupial dry homologue is expected to be necessary for testis determination, but confirmation of this hypothesis *will* require functional studies.

Students didn't discuss the future needs their research suggested, perhaps one reason why overuse of *will* is not as noticeable in the Results/Discussion sections as it is in Introductions and Methods sections.

The most common function for *will* in the professional corpus was to indicate inductive reasoning, an accepted or logical truth. This was especially true in the Discussion section, where *will* was used to speak about the results of the experiment and what could now be expected or understood. However, in many occasions where *will* was used, the writers also used hedging to qualify claims, with words like *suggests* and *seems likely*.

This suggests that an antireflux operation *will* not inevitably lead to regression or resolution of Barrett's epithelium, and may not reduce...

Moreover, this fermentation of fibre *will* have effects on colonic ph.

We suggest that the presence of acetylcholine alone *will* not provide an adequate stimulus for the sweat response in the event of nerve degeneration.

Overall, it seems likely that the existence, speed and intensity--correlated behaviour of qpos *will* provide severe constraints on models.

While most occurrences of *will* in the professional corpus functioned inductively, this function was rarer in the student corpus. For the most part students did not use *will* in this way. In those circumstances where students would want to speak about the what logically followed, from their findings, for example, students were less likely to use *will*, the accepted form in the professional corpus, and much more likely to use *can* in an ambiguous use suggesting possibility and ability. Students may possibly be unaware or unable to use other hedging forms and thus rely on the modal *can* when they want to qualify a claim. A final example demonstrates this type of student use of *can*, for a situation in which *will* may be more appropriate:

All electrons are negatively charged, so it might seem surprising that two like charges *can* attract.

From the above examples we see professionals are using *will* in ways students do not. Professionals discuss the future needs in their field with *will* and use *will* in inductive statements.

Summary: Professional-Student Differences in Modal Use

In the quantitative findings we see students generally use more modals than the professionals. They overuse modals in all of the sections of the research article, although the differences seem more pronounced in the Methods sections, where modals are so comparatively rare in the professional corpus.

Differences in use are most pronounced in a few modal auxiliaries, *may*, *can* and *will*. *May* is the only modal auxiliary the students use less than the professionals, to a great degree. Although both groups seem to use *may* for the same function, extrinsically to make an epistemic statement about a proposition, students use this function less often, especially in the Introduction section. These findings could imply that students are not expressing the areas of uncertainty in their field to the same degree or in the same manner as their professional counterparts.

Can occurs far more often in the student corpus than in the professional. For students, it not only functions to express intrinsic ability, the most common function in the professional corpus, but also to establish the legitimate way of viewing or understanding a proposition, which seems in some situations the students' way of expressing epistemic meaning. *Can* is often used ambiguously, especially in situations not common in professional use. The student corpus also had greater occurrences of *can* used to relate mathematical terms and formulas than professionals, especially in Methods sections, a result of the different subject matters covered in both corpora. Thus we see students using *can* in ways not as common in the professional corpus: to express legitimacy, as a sort of author comment, in ambiguous meanings, as well as to convey mathematical and scientific

formulas.

Will does not serve to express the inductive meaning in the student corpus as often as it does in professional scientific writing. Rather, students most often use *will* to express their intentions for the immediate future within the paper or within the experiment they are carrying out, a function rarely used by the professionals. Students do overuse *will* inductively in Methods sections in the fields of engineering and physics.

From these findings we can see striking differences in the frequencies and functions of use of modal auxiliaries by the two groups. The implications of these findings will be discussed in the next chapter.

CHAPTER 5. CONCLUSION

This study found some important differences in the ways learners of English and professional scientists use modal verbs in scientific research articles, differences in degree of use as well as differences in the types of modal used and the purposes or functions for which they are used for. This information is helpful for students, their teachers and materials writers. The study's results have additional application for further research. This chapter presents the major findings of the study, limitations of the study and implications for further research and applications of the findings for students and educators.

Major Findings

The research question which drove this study was divided into four, closely-related questions. The first question, which addressed the issue of the differences in use of modal verbs in scientific research articles by NS professionals and advanced international graduate students with respect to the overall frequency of modal verbs used, was answered by counts of all modal verb occurrences in both corpora. The results of these counts indicate that students used modal verbs twice as frequently as professional writers did.

Results of the second question concerning the frequencies in different research article sections determined that while professionals and students both used more modals in Results and Discussion sections than in Methods sections, students

overused modal verbs in every section, and overused them to the highest degree in the Methods sections.

The third question was concerned with the differences in use of individual modal verbs. Three modal verbs in particular exemplified the differences in student use. *Can* and *will* were the most generally overused modals by students and *may* was the one modal consistently underused by students. *May* was underused most in Introductions and Methods sections. *Can* was overused most significantly in the Methods and Results/Discussion sections, while *will* was overused most in the Introduction and Methods sections.

The fourth question addressed the issue of functional use of the modal verbs. In their use of *may*, students and professionals performed the same functions with the verb, expressing possibility; students, though, used the modal for this function less often than their professional counterparts, suggesting students write more direct, unmitigated statements than are normally found in professional writing.

Use of *can* and *will* differed not only by degree but also by the functions these modal verbs fulfilled in the two corpora. Students tended to overuse *can* for many different functions including functions professionals more often performed with *may* and *will*. Students used *can* to suggest possibility and to indicate a legitimate understanding of a proposition in situations where professionals would use *may* to make an extrinsic comment on the certainty of a proposition. Students often used *can* in ambiguous ways, and *can* was also heavily used in students' engineering and physics texts in discussions of mathematical formulas, especially in Methods sections. In part, this overuse of *can* may be due to differences in topics.

Will was overused by students in Introductions to indicate volition, often co-occurring with first-person pronouns, and to give an overview of the study. Students also used *will* in Introductions to discuss the implications their study would have, a function professionals used in Results/Discussion sections. Students also overused *will* in the Methods sections to express induction in engineering problems and formulas. This again suggests the influence differing writing topics had on the study's results. Students did not use *will* for some functions which were common in the professional corpus, including indicating induction and discussing the future needs of the field.

Finally, we see students overusing modals throughout the corpus and in each research article section. They especially overuse *can* and *will*, while underusing *may* and they often use these verbs to perform functions for which the professionals do not use these modals.

Of course, we must be careful not to over-generalize these findings. The small student corpus, which was the result of a limited number of student papers and student papers with fewer words than most professional articles, causes the findings to be less generalizable, and further study with a larger student corpus is necessary to determine if the findings of this study are applicable to most international student writers. As mentioned earlier, differences in writing topics also limited this study. Further research needs to be done to determine how much affect different subject areas have on the modal verb use by students. Nevertheless, this study has important implications for those concerned with the scientific writing of non-native speakers.

Implications for Materials Development and Current Instruction

The findings presented in this study will be especially useful to ESL instructors teaching advanced research writing and to material writers preparing EAP and ESP textbooks. The varied findings of this study provide ample possibilities for additional instruction. Not only would non-native speakers benefit from such teaching, but native speaking students would also likely be helped by learning the appropriate contexts for modal verbs, especially in technical communication.

Teachers can help students understand the different, less prominent, role modal verbs play in research writing compared with more informal, subjective types of writing. This might be done through exposure to conversation texts, or to more informal personal writing. Students could be asked to compare the ways modal verbs are used in these situations with the ways the forms are used in research writing. Students could be directed to pay special attention to the common use of personal pronouns and human agents with modal verbs in conversation and informal writing which does not often occur in academic writing. Academic writing textbooks could also address the different qualities of conversation and academic writing, and highlight specifically the use of modal verbs. See Figure 5.1 for an example of a class exercise addressing differences in modal use between registers. In this example, students used are asked to look at an example of a conversation and an example of a research article discussion section. They are instructed to pick out the modal verbs and their noun subjects and determine what functions the modals are playing in both examples. Although they have extremely limited data, students are

able to see that modal verbs occur more often in conversation with personal pronouns as the subjects, and they are used more often to express ability and intention, while in research writing the modal verbs are used to express possibility. This exercise could be extended with more examples which would make the findings more generalizable for the students.

Students should also be made more aware of the functional appropriateness of modal verbs in Introductions and Discussion sections rather than in Methods sections. Exercises which ask students to count the numbers of modal verbs in each section of a research article and to make comparisons between the sections and then to use these findings to make generalizations about the functions modal verbs play in professional research articles would be helpful. See Figure 5.2 for an example of such an exercise. In this example, the students are asked to compare the findings of this study, which show that modal verbs are used least often in Methods sections, with writing from their own disciplines. They are to count all occurrences of modal verbs in each section of a research article from their field and determine where modal verbs are used most frequently. They are then asked to write a comment statement for a Discussion section expressing why modal verbs are used differently in different research article sections.

Because students are more likely to use modal verbs intrinsically, often with a human agent, rather than extrinsically, to express doubt and uncertainty, students need to see how statements are qualified throughout professional research articles, with *may* as well as with other hedging forms, including modal adverbs, adjectives and lexical verbs. In Figure 5.3 students are asked to write all occurrences of

"hedges" which they find in an example research article from their field. They are asked to summarize their findings and suggest why hedges are used in certain situations.

Textbooks need to address the various hedging forms, and teachers should spend class time discussing the appropriate contexts for hedging forms, including the most common hedging forms, modal verbs.

Students should also be taught to recognize when they use these forms for unnatural functions, such as *can* to indicate legitimacy. A concordancing exercise exposing students to the most common uses of *can* and *may* in research writing would help students learn the appropriate functions inductively. In Figure 5.4 students use a concordancer, Monoconc, to retrieve all the occurrences of *can* and *may* in a limited corpus of research articles. Students are to examine the subjects used with the verbs and the functions these verbs play in the writing. They are then expected to compare the uses of these two modal verbs to notice differences in their meanings and uses. This exercise, along with the others previously explained, will allow students to inductively discover some of the natural, standard uses and functions of modal verbs in research writing.

Another way teachers might begin to address these issues could be through student editing exercises in which they are asked to find and correct specific uses of modal verbs in their writing which are unnatural or uncommon in most professional scientific writing. Textbooks could also give examples of overuse and unnatural use of words like *can* or *will*, helping students to be aware of potential misuse of such words and to recognize what they can do to avoid inappropriate use of the forms.

Modal Verbs. Comparison of registers.

You are going to look at two different registers, or types of communication, and compare how modal verbs are used in each register. The first register is conversation, and the second is research article writing. Work with a partner to analyze the two registers.

A. CONVERSATION
Below is a SMALL piece of conversation between three people. Do the following analysis:

1. Circle all the occurrences of MODAL VERBS: can, could, will, would, may, might, must, shall and should.
2. Circle the noun subjects of those verbs.
3. List the subjects, verbs, and the functions of the modal verbs in the space below.

«Rand» Well don't write it all on Friday, we may have something we want to do.

«Mandy» I can pick you up if I'm not doing anything... Not that I wouldn't pick you up 'cause I was doing anything but I mean if I don't have to go to school or anything.

«Elizabeth» Do you have plans on Friday night ten fifteen p.m.

«Mandy» Forty five.

«Rand» Oh hey, you can pick her up Mandy.

«Elizabeth» Mom!

«Mandy» Okay I will pick you up.

Subjects:	Modal Verbs:	Functions of the verbs
we	may	show possibility

B. RESEARCH WRITING
Follow the same steps of analysis for this portion of a Discussion Section.

Geobacter metallireducens (formerly, strain Os-15) and *Shewanella putrefaciens* are the only well characterized organisms known to couple the oxidation of fermentation acids and/or hydrogen to the reduction of Fe(II). The characteristic Pfla data gave no indication that either of these organisms was present in the mud or the concretions. This, together with the *Desulfovibrio* enrichment, led us to investigate the possibility that *Srb* might be involved in Fe(II) reduction. In aquatic sediments, H₂ is likely to be the most important electron donor for *Desulfovibrio*. A pure culture of *Desulfovibrio* desulfuricans readily reduced Fe(II), whereas *Desulfovibrio* postgatei and *Desulfovibrio curvatus* did not (fig. 2). The reduction of Fe(II) in the presence of *D. desulfuricans* could not be the result of indirect reduction of Fe(II) by sulphide because the initial sulphate concentration of only 320m could provide enough sulphide to reduce less than 5% of the Fe(II) that was reduced. Furthermore, all of the sulphate was reduced within the first day of the experiment whereas most of the Fe(II) reduction was after the first day.

Subject	Modal verbs	Function
Reduction	could	ability/possibility

What differences do you notice between the two registers with regard to:

- the number of modal verbs
- the types of modal verbs
- the types of subjects
- and the functions that the modal verbs play?

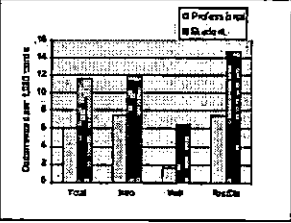
How could these findings effect your own research writing?

Make some comments about your findings here:

Figure 5.1 Worksheet comparing modal use in registers.

Modal Verbs. Comparison of research article sections.

Below is a figure representing the findings of my own research. I found that modal verbs occurred more often in introductions and in results/discussion sections than in methods sections. (You can also see that students always overused modal verbs!)



Section	Professors	Students
Total	10	12
Intro	10	10
Methods	6	2
Results/Discussion	14	14

Look at one of your own example research article and count the number of modal verbs in each section to see if the writing there reflects my findings.

What do you find?

Explain how your example supports or refutes these findings.

Why do you think it is true that modal verbs are used less often in methods sections? Give an explanation statement—one that I might use in my discussion section!

Figure 5.2 Worksheet studying modal use in article sections

Modal verbs and other tentative words.

We've talked a little about tentative verbs, suggesting that they are useful in results and discussion sections because they help make a claim sound more reasonable or precise. We've also mentioned that other words can also be tentative.

Here are some possible tentative words. They aren't all verbs!

Verbs: *may, could, might, would appear, suggest, seem,*

Indicate,

Adjectives: *possible, probable, likely*

Adverbs: *possibly, probably, apparently, approximately*

Look at your sample article again. This time find all of the tentative words in the each research article section. My list is just to get you started. See how many more you can find. Based on your findings, which sections of the research article are more tentative? Why do you think this is true?

Introduction	Methods	Results	Discussion
Suggest			

Write a summary of your results and explain why you think this happens.

Figure 5.3 Worksheet examining hedging forms.

Modal Verbs: Analysis of *can* and *may*

These are results from a concordancing software program. The program has found every example of *can* in our sample of research articles.

Analyze the results to discover how *can* is most often used in research articles.

1. Pick 20 examples and list a) the subjects of the verb b) the functions of the verb: to show possibility, ability, legitimacy, mathematical functions, or more than one of these. Discuss with a partner.
2. Make some generalizations about the ways *can* is used in scientific research articles. Use examples from the sample sentences.
3. How are *can* and *may* used differently in research articles?

Concordance 9 (can)

1. stem gains over five times the flexural stiffness, and consequently can withstand over five times the load of the supported stem along the lower two thirds of its
2. ... site of the present study, where can be 30 meters tall and have basal stem diameters 16 cm, and shrubs can be up to 3.8 meters tall with basal stem diameters to 20 cm. Plants can also be found growing in a can
3. flexural stiffness, and consequently withstand over five times the load
4. however, species of *Callistobranchius* can coexist in the laboratory (Ulida 19
5. ... best use, suitability of host status can be affected by secondary metabol
6. loss and A. incarnate-reared queens can be directly or indirectly exclud
7. is, so that no overall seasonal pattern can be detected in the moist meadow
8. ... its that growing season length can be a limiting factor in alpine tur
9. ... ong the variables. Path analysis can test a particular hypothesis in
10. 73; Craig 1977). But explanations can be dismissed as a mechanism e

10 sentences Original text only Print nothing on

Figure 5.4 Example of a concordancing exercise.

Implications for Further Research

The direct instructional applications of this research suggest the importance of further research into the use of modal verbs. The research questions in this study have not been addressed to any significant extent in the past, and thus the needs of future research are extensive. As was mentioned earlier, similar studies be done with a larger learner corpus and with a corpus of learner writing which more closely matches the subjects in the professional corpus. Further studies must also address the questions of topical influence on modal verb use. Research should also be done comparing the lexical verbs and the agents co-occurring with the modal verbs. The initial findings indicate students tend to use more personal pronouns with modal forms, suggesting a more conversational or involved tone achieved at least in part through modal verb use, but again, further work must be done. Related studies concerning the predominance of active and passive voice in modal verb use would also be helpful. Additionally, many applied linguists have studied the role of other modal forms, modal adverbs, adjectives, and lexical verbs and nouns that imply modality, including hedging, and such research should also be done on the learner use of these forms.

Conclusion

This study used two corpora to examine student use of modal verbs in research writing and found that students generally overuse most forms throughout research articles and in every research article section when compared to the

professional corpus. Only *may* was significantly underused. Students were also found to use modals, especially *can* and *will*, for different purposes than professionals.

Of course the findings must be considered in the light of the study's limitations, its size of the student corpus and the differences in research topics, which may affect the modals used. However, the findings do give us important information about learner use of these important forms.

Students should be served by the information in this study, which will allow teachers and textbook writers to better inform them about the appropriateness of modal verbs in scientific writing and the nuances of functions fulfilled by these modal verbs in competent research writing.

APPENDIX A: CONSENT FORM

This appendix contains a copy of the consent form the subjects were asked to sign.
The form appears on the next page.

To: 101d students: Fall '99
From: Jenny Hykes, MA candidate in English
Re: consent to participate in a master's thesis project

I am a graduate student in Teaching English as a Second Language/Applied Linguistics. I am doing my thesis research on how international students write research articles. My research should help international students improve their scientific writing.

For my research I am asking for your permission to copy the rough draft of your research article that you hand in to your English 101d instructor.

You will not need to make any extra copies or do any additional writing. Your instructor and I will not discuss your paper or my analysis of your writing, and your grade will not be affected in any way by my research. The data I receive from your writing will be reported without your name.

Your participation in this project is completely voluntary, and you can end your participation at any time.

If you are willing to participate in this project, please sign the attached form, and I will automatically make a copy of the rough draft of your article when you hand it in to your instructor.

Thank you for your help.

Jenny Hykes

I agree to allow Jenny Hykes to use a copy of my research article in her thesis research on international students' writing.

I understand I can withdraw my participation in this project at any time.

Name _____
(printed) (signed)

Date _____

English 101D section/instructor _____

Major _____

Years attending ISU _____

Years studying in the U.S. or another English speaking country _____

Native language _____

APPENDIX B: HUMAN SUBJECTS APPROVAL

This appendix contains a copy of the form "Information for Review of Research Involving Human Subjects" used at Iowa State University. The form shows that approval for the study was granted by the Human Subjects Committee. The form appears on the next three pages.

Information for Review of Research Involving Human Subjects

Iowa State University

(Please type and use the attached instructions for completing this form)

1. Title of Project A comparison of the use of modal verbs in research articles by NS professionals and NNS graduate students.

2. I agree to provide the proper surveillance of this project to insure that the rights and welfare of the human subjects are protected. I will report any adverse reactions to the committee. Additions to or changes in research procedures after the project has been approved will be submitted to the committee for review. I agree to request renewal of approval for any project continuing more than one year.

Jenny Hykes

Typed name of principal investigator

10.1.99

Date

English

Department

206 Ross Hall

Campus address

296-2072

Phone number to report results

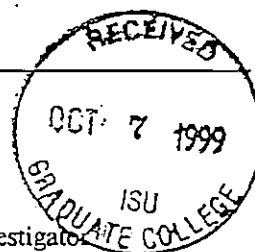
Signatures of other investigators

Date

Relationship to principal investigator

10/5/99

Major professor



4. Principal investigator(s) (check all that apply)

☐ Faculty

☐ Staff

☒ Graduate student

☐ Undergraduate student

5. Project (check all that apply)

☐ Research

☒ Thesis or dissertation

☐ Class project

☐ Independent Study (490, 590, Honors project)

6. Number of subjects (complete all that apply)

adults, non-students: _____

minors under 14: _____

minors 14 - 17: _____

ISU students: 68

other

(explain): _____

7. Brief description of proposed research involving human subjects: (See instructions, item 7. Use an additional page if needed.)

The purpose of this research project is to compare how native speakers and international graduate students use modal verbs in their writing. I will take samples of writing from students in Fall 1999 English 101d courses. The students will not have to do any extra work or make any extra copies of their writing. After obtaining a copy of their writing from their instructor, (with their permission), I will count the number of modal verbs used and interpret the functions these modal verbs are playing in their writing and compare these findings with the use of modal verbs by native speaking professionals. Any reporting of the data will be done anonymously, and subject's participation will be completely voluntary.

(Please do not send research, thesis, or dissertation proposals.)

8. Informed
Consent:

☒ Signed informed consent will be obtained. (Attach a copy of your form.)

9. Confidentiality of Data: Describe below the methods you will use to ensure the confidentiality of data obtained. (See instructions, item 9.)

No names will be published with the data I obtain. I will not discuss the subjects with their instructors.

10. What risks or discomfort will be part of the study? Will subjects in the research be placed at risk or incur discomfort? Describe any risks to the subjects and precautions that will be taken to minimize them. (The concept of risk goes beyond physical risk and includes risks to subjects' dignity and self-respect as well as psychological or emotional risk. See instructions, item 10.)

No risks or discomfort. Participation is completely voluntary. _____

11. CHECK ALL of the following that apply to your research:

- ☐ A. Medical clearance necessary before subjects can participate
☐ B. Administration of substances (foods, drugs, etc.) to subjects
☐ C. Physical exercise or conditioning for subjects
☐ D. Samples (blood, tissue, etc.) from subjects
☐ E. Administration of infectious agents or recombinant DNA
☐ F. Deception of subjects
☐ G. Subjects under 14 years of age and/or ☐ Subjects 14 - 17 years of age
☐ H. Subjects in institutions (nursing homes, prisons, etc.)
☐ I. Research must be approved by another institution or agency (Attach letters of approval)

If you checked any of the items in 11, please complete the following in the space below (include any attachments):

Items A-E Describe the procedures and note the proposed safety precautions.

Items D-E The principal investigator should send a copy of this form to Environmental Health and Safety, 118 Agronomy Lab for review.

Item F Describe how subjects will be deceived; justify the deception; indicate the debriefing procedure, including the timing and information to be presented to subjects.

Item G For subjects under the age of 14, indicate how informed consent will be obtained from parents or legally authorized representatives as well as from subjects.

Items H-I Specify the agency or institution that must approve the project. If subjects in any outside agency or institution are involved, approval must be obtained prior to beginning the research, and the letter of approval should be filed.

Last name of Principal Investigator Hykes**Checklist for Attachments and Time Schedule**

The following are attached (please check):

12. ☒ Letter or written statement to subjects indicating clearly:
- a) the purpose of the research
 - b) the use of any identifier codes (names, #'s), how they will be used, and when they will be removed (see item 17)
 - c) an estimate of time needed for participation in the research
 - d) if applicable, the location of the research activity
 - e) how you will ensure confidentiality
 - f) in a longitudinal study, when and how you will contact subjects later
 - g) that participation is voluntary; nonparticipation will not affect evaluations of the subject
13. ☐ Signed consent form (if applicable)
14. ☐ Letter of approval for research from cooperating organizations or institutions (if applicable)
15. ☐ Data-gathering instruments

16. Anticipated dates for contact with subjects:

First contact

October 11, 1999

Month/Day/Year

Last contact

December 10, 1999

Month/Day/Year

17. If applicable: anticipated date that identifiers will be removed from completed survey instruments and/or audio or visual tapes will be erased:

December 10, 1999

Month/Day/Year

18. Signature of Departmental Executive
Office

Date

Department or Administrative Unit

10-6-99English

19. Decision of the University Human Subjects Review Committee:

☒ Project approved☐ Project not approved☐ No action required

Name of Human Subjects in Research Committee Chair

Date

Signature of Committee Chair

Patricia M. Keith10/14/99

3

WORKS CITED

- Adams Smith, D.E. (1984). Medical discourse: Aspects of author's comment. *ESP Journal* 3, 25-36.
- Atkinson, D., Biber, D., Finegan, E. (1994). ARCHER and its challenges: Compiling and exploring: A Representative Corpus of Historical English Registers. In Fries et al. (eds.), 1-14.
- Biber, D. (1998). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D., Conrad, S., Reppen, R. (1998). *Corpus Linguistics Investigating language structure and use*. Cambridge: Cambridge University Press.
- Biber, D., Johnansson, S. Leech, G., Conrad, S. and Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*. London: Longman.
- Butler, C. (1990). Qualifications in Science: Modal Meanings in Scientific Texts. *The Writing Scholar*. Ed. Walter Nash. London: Sage Publications.
- Coates, J. (1983). *The Semantics of Modal Auxiliaries*. Croom Helm: London.
- Conrad, S. (1996). *Academic Discourse in Two Disciplines: Professional Writing and Student Development in Biology and History*. Unpublished Ph.D. dissertation. Flagstaff, AZ: Northern Arizona University.
- Conrad, S. (1999). The importance of corpus-based research for language teachers. *System* 27:1-18.
- Cook, W. A. S. J. (1978). Semantic Structure of the English Modals. *TESOL Quarterly* 12 (1), 5-16.
- Flowerdew, L. (1997). Interpersonal Strategies: Investigating Interlanguage Corpora. *RELC Journal* 28, 72-88.
- Flowerdew, L. (1998). Integrating 'Expert' and 'Interlanguage' Computer Corpora Findings on Causality: Discoveries for Teachers and Students. *English for Specific Purposes* 17 (4), 329-345.

- Gibbs, Dorothy. (1990). Second Language Acquisition of the English Modal Auxiliaries can, could, may and might. *Applied Linguistics* 11(3), 297-313.
- Granger, S. (1998). The computer learner corpus: a versatile new source of data for SLA research. In Granger, S. (ed.) *Learner English on Computer*. Longman, New York, 3-18.
- Halliday, M.K. (1976). *Halliday: System and Function in Language*. Kress, G.R., Ed. Oxford University Press: London.
- Hinkel, Eli. (1995). The Use of Modal Verbs as a Reflection of Cultural Values. *TESOL Quarterly* 29 (2), 324-341
- Holmes, J. (1982). Expressing Doubt and certainty in English. *RELC Journal*. 13(2), 9-28.
- Holmes, J. (1983). Speaking English with the appropriate degree of conviction in C. Brumfit (ed.) *Learning and Teaching Languages for Communication: Applied Linguistics Perspectives*. London: CILT.
- Holmes, J. (1988). Doubt and Certainty in ESL Textbooks. *Applied Linguistics* 9, 21-43.
- Hu, Z., Brown, D., Brown, L. (1982). Some Linguistic Differences in the Written English of Chinese and American students. *Language Learning and Communication*. 1(1), 39-49.
- Hyland, K. (1994). Hedging in Academic Writing and EAP Textbooks. *English for Specific Purposes* 13(3), 239-256.
- Hyland, K. (1996). Nurturing Hedges in the ESP Curriculum. *System* 24 (4): 477-490.
- Hyland, K. and Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing* 6 2: 183-205
- Leech, G. and Coates, J. (1980). Semantic Indeterminacy and the Modals. *Studies in English Linguistics*. (ed. Greenbaum, Leech and Svartvik.), 79-90.
- Milton, J. and Tsang, E. (1993). A corpus-based study of logical connectors in EFL students' writing. In Pemberton, R. and Tsang, E. (eds) *Studies in Lexis*. Hong Kong University of Science and Technology, Hong Kong, 215-46.

- Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics* 10, 1-35.
- Palmer, F.R. (1979). *Modality and the English Modals*. Longman: London.
- Reid, J. (1992). A computer text analysis of four cohesion devices in English discourse by native and nonnative writers. *Journal of Second Language Writing*, 1 (2), 79-107.
- Salager-Meyer, F. (1994). Hedges and Textual Communicative Function in Medical English Written Discourse. *English for Specific Purposes* 13 (2), 149-170.
- Skelton, J. (1997). The Representation of Truth in Academic Medical Writing. *Applied Linguistics* 18 (2), 121-140.
- Stubbs, M. (1994). Grammar, Text, and Ideology: Computer-assisted Methods in the Linguistics of Representation. *Applied Linguistics* 15 (2), 201-223.
- Swales, J. (1990). *Genre Analysis*. Cambridge: Cambridge University Press.
- Thompson, G. and Yiyun, Y. (1991). Evaluation in the Reporting Verbs Used in Academic Papers. *Applied Linguistics* 12 (4), 365-382.
- Varttala, Teppo. (1998). Remarks on the Communicative Functions of Hedging in Popular Scientific and Specialist Research Articles in Medicine. *English for Specific Purposes*. 177-199.

ACKNOWLEDGEMENTS

Two years ago, I didn't know what a modal verb was. I'd never seen a research article. I didn't know how to teach anyone anything. I had no idea what international students even needed to learn.

I thank all of my professors in the TESL program who have taught me so much and enriched my life in so many ways during these two years. Thank you, Barbara Mack, for challenging me in many ways. Roberta Vann, you have transformed my teaching with your patient and thorough teaching and guidance.

Susan Conrad, this thesis is a small piece of evidence of the great impact you have had in my work. Your teaching, gentle directing, and thoughtful editing are all displayed in these pages. More than all of these, however, the example of your own commitment to excellence in research, in teaching and in all you do have helped to guide my vision of how I want to develop professionally.

My teachers have been more than classroom teachers, and my classmates, more than classroom "mates." These friends have challenged me to think more clearly and work harder in every project. They also have a part in this work.

Most of all, my God, Jesus Christ, deserves my thanks. You have blessed my life with so many who reflect you in their work, teaching and friendship. You gave me a dream two years ago to love and serve people from other cultures. You have opened up areas of knowledge which have brought surprising fulfillment. These are small examples of the unending work of your redemptive love in one life. Thank you.